BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

PROCEEDINGS II OF THE 29TH STUDENT EEICT 2023

ELECTRICAL ENGINEERING, INFORMATION SCIENCE,
AND COMMUNICATION TECHNOLOGIES

SELECTED PAPERS

April 25, 2023
Brno, Czech Republic

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Title	Proceedings II 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 978-80-214-6154-3

ISSN 2788-1334

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

SPONSORS

Diamond tier

Honeywell

Golden tier

ABB

onsemi

ThermoFisher
SCIENTIFIC

SPONSORS

Silver tier

ASICentrum — A COMPANY OF THE SWATCH GROUP

Garrett — ADVANCING MOTION

FORVIA — HELLA

NXP

TERADYNE

Valeo

Honorary tier

e·on

SKODA

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

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. These Proceedings comprise 60 award winning full papers, all selected by the conference's evaluation boards.

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

Detection of Gunshots from Small Arms

Ondřej Nesvadba
Brno University of Technology
Department of Radio Electronics
Brno, Czech Republic
xnesva08@vutbr.cz

***Abstract*—This paper deals with acoustic gunshot detection from small arms primarily for use in urban areas. Key part of the paper is an examination of typical features of gunshot signals. Based on the short-term features detection algorithms are proposed in the time and frequency domains. Training and evaluation of the algorithms was performed using real gunshots, non-gunshots as well as synthetic gunshots. The non-gunshots set contains impulsive acoustic events such as dog barking, glass breaking, car horn, etc.**

***Keywords*—detection, detection algorithms, frequency domain, gunshot, gunshot parameters, small arms, time domain**

I. Introduction

This paper deals with the issue of automatic gunshot detection, especially in urban areas with high crime rates. The goal is to propose a comprehensive system that recognizes a gunshot from other impulsive background sounds, such as breaking glass, car horn or barking dog. The main motivation for this study is to develop a system that reliably detects gunshots and immediately sends an alarm message to the responsible authorities.

In general, automatic gunshot detection can be deployed in three principle different scenarios according to the surrounding environment, namely: detection in open space (i.e. nature), detection in urban areas and detection in indoor environments. The purpose of detection in open space is often to protect against poachers [1] [2] or for military use [3], while detection in urban areas and in public indoor spaces is mainly aimed at increasing security. A presentation of gunshot detection technologies effective in urban environments can be found in [4]. A surveillance system for detection of gunshots in indoor environment is described, for example, in [5]. A comparison of successful detection algorithms can be found in the proceedings of the "Detection and Classification of Acoustic Scenes and Events" competition, which was focused on sound detection and the target events were gunshots [6]. Some simple gunshot features, which need relatively low computational time, are proposed in [7].

Focus of this paper is to identify real gunshots in urban environments. For that, analysis of signal waveforms and knowledge of basic gunshot characteristics are needed in order to distinguish real gunshots from other highly impulsive signals that sound similar. From the practical point of view, it is important to eliminate false alarms [8]. To achieve high reliability, fake synthetic gunshots were also examined and compared to the real ones.

II. Gunshot Characteristics

Gunshot itself is a very complex physical phenomenon. When the trigger is pressed, multiple signals are propagating through the surrounding. Two most important of these signals are the muzzle blast and the shock wave. A muzzle blast is an audio signal that propagates directly to the microphone. The best way to record a muzzle blast is when the microphone is on the axis of the gunshot. On the other hand, if the microphone is behind the weapon, the muzzle blast is attenuated. A shock wave can be created when the shot is supersonic. The way of shock wave is conical in the direction of the gunshot, so the microphone placed behind the weapon does not pick up the signal at all.

To record a gunshot, only one microphone is necessary. The ideal position is, as already mentioned, right in the axis of the gunshot. However, the more microphones are used, the more information can be obtained. Using four microphones, we can deduce the speed and direction of the gunshot and even the location of the shooter, which can be very helpful in finding the criminal. This issue is not included in this study, but the identification of the shooter's position is in view as a logical extension of this work in the future.

The gunshot signal has a characteristic waveform shown in Fig. 1. It has its typical N-shaped part. The whole gunshot lasts about ten milliseconds. The gunshot is highly impulsive and its duration is very short with few significant extremes at the beginning. This distinguishes it from other sounds in the time domain. They are usually not that short with the first three to five extremes being so significant.

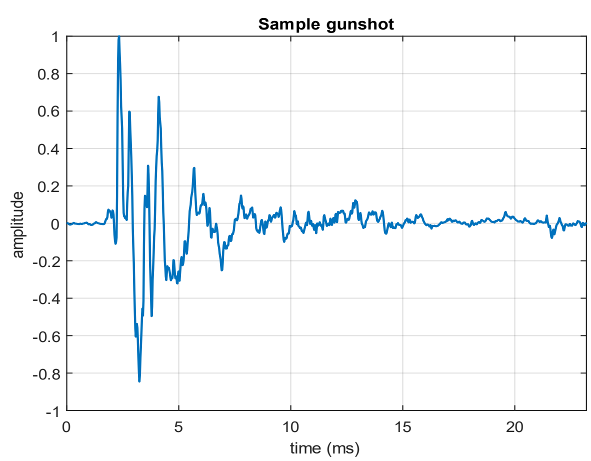


Fig. 1. Example of a typical gunshot waveform.

III. Methods for Gunshot Detection

In this section, a few methods used commonly for gunshot detection and used in this study as well will be theoretically introduced. We can analyze signals in time domain or frequency domain. This is the same for gunshot detection. Even algorithms proposed further in this paper are both in time domain as well as in frequency domain.

A. *Correlation*

Correlation is probably the most important and most commonly used method in time domain. It expresses a similarity of two examined signals. In [9], correlation was stated as the most successful method for gunshot detection. For the proposed algorithm, we need a sample signal of gunshot. Other examined signals are afterwards correlated with the sample signal. As a sample signal, self-recorded gunshot from [10] was chosen. Correlation is defined as

$$R_{fg}(t) = \int_{-\infty}^{\infty} f^*(\tau) g(t+\tau) d\tau \qquad (1)$$

where f and g are correlated signals and τ is mutual time shift.

B. *Fourier Transform*

To investigate the gunshot signals in the frequency domain, the spectrum of the signals was calculated by the Fourier transform defined as

$$S_{(\omega)} = \int s(\tau) e^{-j\omega\tau} d\tau \qquad (2)$$

where $S(\omega)$ is spectrum of signal $s(t)$. When spectrum is created, further spectral analysis can be done like spectrum energy or spectral entropy. A comparison of the spectra obtained from gunshot and bottle break is shown in Fig. 2.

IV. Detection parameters

Key part of the whole study is extraction of typical values from gunshot and non-gunshot recordings. From these values, parameters of detection algorithms are determined.

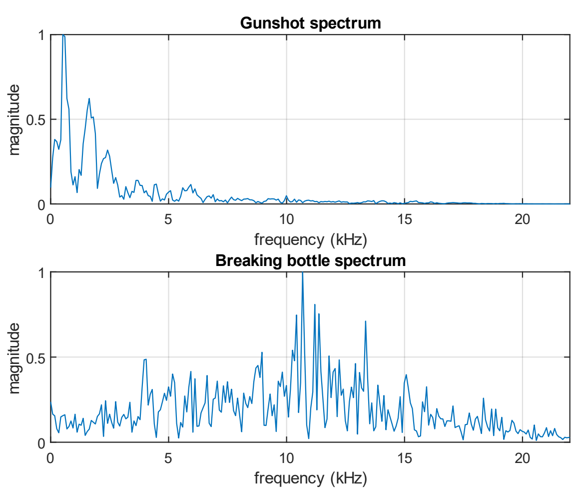


Fig. 2. Example of gunshot and non-gunshot spectrum.

First step was to obtain signals from recordings. That was processed in Matlab. From each recording, two signals were made. Firstly, longer signal of 20 480 samples with the whole gunshot lasting (further called "Sig") and then shorter signal of 512 samples centered around the maximum amplitude of the gunshot (further called "Shot"). To be unified, signals were standardized to the maximum amplitude equal to one. From those unified signals, further described parameters were computed using Matlab.

A. *Average Signal Energy*

First parameter was an average spectral energy. As gunshots are impulsive, the energy is high, but only in short time section. That means average energy, especially of the longer "Sig", is not so high. On the contrary it is low compared to the shorter "Shot" That will be used in one of the algorithms. Other reason for using this parameter is simple differentiation from signals without any impulse.

B. *Correlation with Sample Signal*

Correlation seems to be the most important parameter for the time domain-based algorithm. Signal under test is correlated with sample signal and sum of the correlation is computed. The higher the value is the more similar signals. We apply correlation only for shorter "Shot".

C. *Energy Distribution*

This parameter is related to the average signal energy. Number of samples of the signal that have larger value than 75 % of maximal amplitude is computed. The number should be smaller for gunshots because of short duration of the impulse.

D. *Spectral Energy*

As gunshot is short impulse, the spectrum is not very rich. There are only a few spectral components close to the beginning of whole spectrum. That means spectral energy should be low. That is one of the parameters of frequency domain-based algorithm. Other non-gunshot recordings have richer spectrum therefore greater spectral energy.

E. *Spectral Difference*

Similar to correlation, we use sample signal to this parameter as well. Spectrum of sample signal is created, then spectrum of examined signal is created and finally the difference of those two spectrums is computed. The difference should be very small for a gunshot and on the other hand quite high for non-gunshots.

F. *Sig-to-Shot Ratio*

As mentioned, average signal energy of gunshot is high for shorter "Shot" but quite low for longer "Sig" because of very short duration if the impulse. That means ration between those two values should be high. At least a lot higher than for vast majority of non-gunshot recordings. This parameter will be used as main detection parameter for one of the algorithms.

G. Summary of the Parameters Obtained

From the values obtained from gunshot and non-gunshot signals, parameters for detection algorithms were determined and the summary of all the parameters are shown in Table 1. Each parameter is assigned to algorithm, in which it is used.

V. Detection Algorithms

Three detection algorithms were proposed and tested. One is based on time domain, one is based on frequency domain and the last one is based on ratio between longer and shorter parts of gunshot signal (Sig-to-Shot ratio).

When an unknown recording is tested, at first unified signals “Sig” and “Shot” are created and all parameters are computed. Then the algorithm compares the computed values of examined signal with the reference parameters as stated in the summary Table 1. If the values of examined signal meet all the conditions, the recording is claimed as gunshot. A flowchart of each algorithm is shown in following figures.

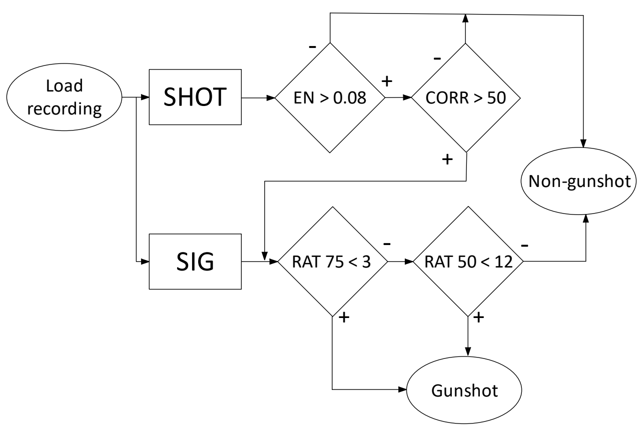


Fig. 3. Flowchart of time domain-based algorithm.

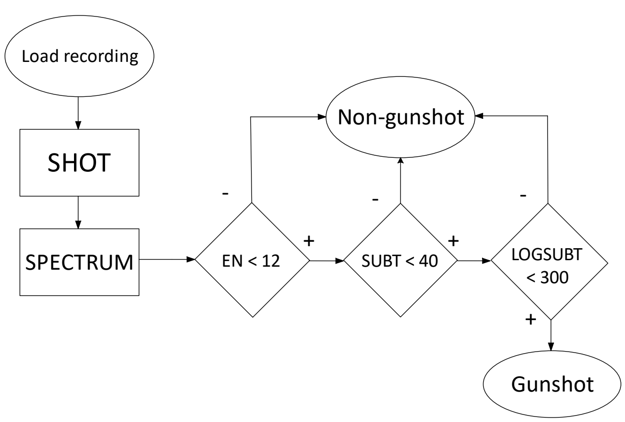


Fig. 4. Flowchart of frequency domain-based algorithm.

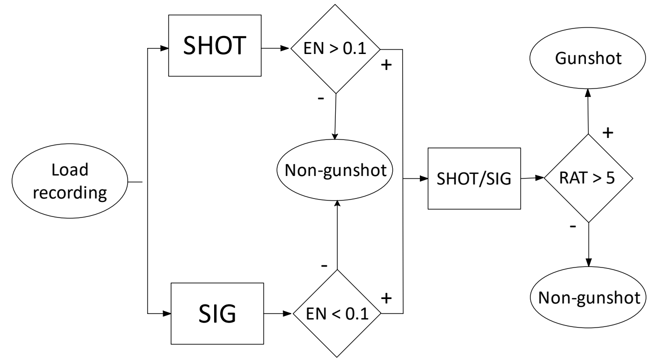


Fig. 5. Flowchart of algorithm based on Sig-to-Shot ratio.

TABLE I. Summary of Parameters.

Summary of parameters		
Algorithm	***Parameter***	***Value***
Time domain	Average signal energy	**> 0.08**
	Correlation	**> 50**
	Energy distribution 75%	**< 3**
	Energy distribution 50%	**< 12**
Frequency domain	Spectral energy	**< 12**
	Spectral difference	**< 40**
	Logarithmic spectral difference	**< 300**
Sig-to-Shot ratio	Average “Sig” energy	**< 0.1**
	Average “Shot” energy	**> 0.1**
	Sig-to-Shot ratio	**> 5**

VI. Obtained Results

The three described algorithms were finally tested on a set of signals covering nine real gunshot signals from 9 mm arm pistol and eight non-gunshot signals including white noise, car horn, dog barking and five glass breaking sounds. Furthermore, two recordings of fake synthetic gunshots, which sound very similar to real gunshots, were tested. All signals used in the experiments were single-channel sounds, sampled at 44.1 kHz and quantized by 16 bits. All recordings were in WAV format.

All proposed algorithms were mostly successful while frequency domain-based algorithm was the best one. All three algorithms successfully detected eight from nine gunshots while vast majority of non-gunshots were ignored. There was one signal, initially considered as real gunshot, which was not detected by any of the algorithms and its waveform was highly extraordinary. It was later decided not to include this signal because it worsened the overall results due to the corrupted recording. Important result is that both fake synthetic gunshots were correctly ignored by all three algorithms. Recordings that sound similarly can be distinguishable by proposed algorithms because their waveform never fully equals to the real gunshots.

TABLE II. SUCCESS RATE OF ALGORITHMS.

Summary of algorithms			
Criterion	***Time domain***	***Frequency domain***	***Sig-to-Shot ratio***
TPR	89 %	89 %	89%
FPR	10 %	0 %	20 %

Table 2 shows success rate of all algorithms. The criterion TPR (true positive rate) is percentage of correctly detected gunshots from all tested gunshots. The criterion FPR (false positive rate) is percentage of incorrectly detected non-gunshots from all tested non-gunshot recordings.

For comparison, two gunshots are displayed below. It is the waveform of a real gunshot (Fig. 6) and a fake synthetic gunshot (Fig. 7), where the difference is clearly visible.

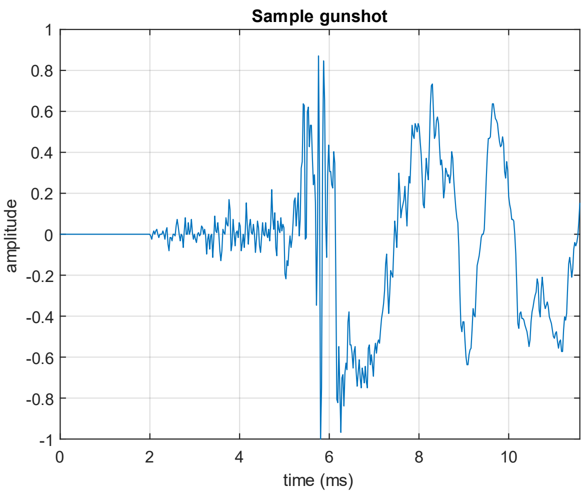


Fig. 6. Real gunshot waveform.

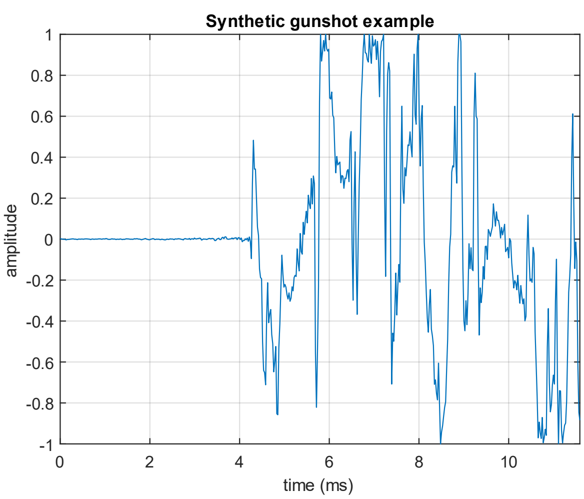


Fig. 7. Fake synthetic gunshot waveform.

VII. CONCLUSION

Three algorithms were proposed to detect of real gunshots while ignoring background noise and other impulsive sounds, as well as fake synthetic gunshots. All algorithms succeed with a success rate of 80 %. By combining these algorithms, a comprehensive detection system can be developed for security purposes in high crime urban areas.

There is also an idea of a mobile application implemented into smartphones as in [11] with detection algorithms that will be able to detect a gunshot and immediately inform the responsible authorities. The more users the application would have, the more information would be obtained. If this application would be used by more users in surrounding of a shooter, the location of the shooter can be easily determined.

In near future, it will be necessary to enlarge our sound database by adding recordings of other types of impulsive sounds and non-stationary noise [12] potentially found in urban environments. Furthermore, it will be useful to mix gunshots with background sounds such as traffic noise, music and human voice.

REFERENCES

[1] G. Kalmar, G. Wittemyer, P. Völgyesi, H. B. Rasmussen, M. Maroti, and A. Ledeczi, "Animal-borne anti-poaching system," in *Proc. Int. Conf. on Mobile Systems, Applications, and Services*, Seoul (South Korea), 2019, pp. 91–102.

[2] M. Sigmund and M. Hrabina, "Efficient feature set developed for acoustic gunshot detection in open space," *Journal Elektronika ir Elektrotechnika*, vol. 27, no.4 pp. 62–68, 2021.

[3] J. A. Mazurek *et al.*, "Boomerang mobile counter shooter detection system," in *Proc. of SPIE – The International Society for Optical Engineering*, vol. 5778 (PART I), Orlando (USA), 2005, pp. 264–282.

[4] J. R. Aguilar, "Gunshot detection systems in civilian law enforcement," *Journal of the Audio Engineering Society*, vol. 63, no. 4, pp. 280–291, 2015.

[5] S. U. Rahman, A. Khan, S. Abbas, F. Alam, and N. Rashid, "Hybrid system for automatic detection of gunshots in indoor environment," *Multimedia Tools and Applications*, vol. 80, pp. 4143–4153, 2021.

[6] *Detection and Classification of Acoustic Scenes and Events* (competition website). [Online]. Available: http://www.cs.tut.fi/sgn/arg/dcase2017/

[7] M. Hrabina and M. Sigmund, "Gunshot recognition using low level features in the time domain," in *Proc. of International Conference Radioelektronika,* Praha, 2018, pp. 1–5.

[8] J. Bajzik, J. Prinosil, R. Jarina, and J. Mekyska, "Independent channel residual convolutional network for gunshot detection," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, pp. 950–958, 2022.

[9] A. Chacon-Rodriguez, P. Julian, L. Castro, P. Alvarado, and N. Hernández, "Evaluation of gunshot detection algorithms," *IEEE Transactions on Circuits and Systems*, vol. 58, no. 2, pp. 363–373, 2010.

[10] M. Hrabina and M. Sigmund, "Acoustical detection of gunshots," in *Proc. of International Conference Radioelektronika,* Pardubice, 2015, pp. 150–153.

[11] D. Welsh and N. Roy, "Smartphone-based mobile gunshot detection", in *Proc. of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops*, Kona (USA), 2017, pp. 244–249.

[12] P. Zelinka and M. Sigmund, "Hierarchical classification tree modeling of nonstationary noise for robust speech recognition," *Information Technology and Control*, vol. 39, no. 3, pp. 202–210, 2010.

Compression of Vehicle-Driving Data by Means of Orthogonal Bases

1st Vít Šimeček
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
230187@vut.cz

2nd Ondrej Mihálik
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
ondrej.mihalik1@vut.cz

***Abstract*—The paper deals with application of orthogonal bases in signal approximation with the aim of data compression in a vehicle driving simulator. Three different bases are tested: Discrete Fourier Basis, Discrete Cosine Basis, and Slepian Basis. Quality of signal approximation error is assessed in terms of global squared errors. Thus obtained numerical results suggest that Slepian Basis affords the sparsest representation of signals tested in this study. Therefore, a considerable reduction of required memory can be accomplished.**

Keywords—signal, approximation, compression, Slepian sequences, DPSS.

I. Introduction

Modern flight or vehicle-driving simulators are capable of simulating complex scenarios under constant or varying conditions, and with a large number of input parameters. The human operator reacts to various stimuli, and, in turn, responds by input devices such as a steering wheel or pedals. Detailed analysis of such human-machine feedback loop is intimately connected with the need of collecting and storing large quantities of data in an efficient way.

This work aims at achieving a sparse representation of real data from a vehicle driving simulator. The simulator records discrete-time signals; therefore, this work will restrict the discussion and analysis to discrete (sampled) time domain.

The paper is divided as follows. Section II. introduces preliminaries necessary for subsequent sections, such as notation, mathematical tools, and introduction to the Slepian Basis. Section III. presents comparison of Slepian Basis with other bases using the real measured signal.

II. Discrete Prolate Spheroidal Sequencies

Discrete Prolate Spheroidal Sequencies (DPSS), also referred to as Slepian Sequencies, constitute the discrete analogy to continuous-time Prolate Spheroidal Wave Functions (PSWF), also called Slepian Functions.

A. Mathematical Background

Generally, the Slepian Sequencies have a number of interesting properties; we will list only those which are vital for our analysis:

1) Bandlimited signals: A signal $f(t)$ is referred to as bandlimited, if its Fourier Transform $F(\omega)$ is of the form

$$F(\omega) = 0 \Leftrightarrow |\omega| > 2\pi W. \tag{1}$$

Here the W denotes the bandlimit of the signal. Note that every discrete sampled signal is bandlimited due to the Shannon-Nyquist theorem associated with reconstruction formula which employs bandlimited sinc functions [1].

If not mentioned explicitly, all transforms, operations, signals etc. in this work are discrete.

2) The Norm of a vector in ℓ^2 space

$$\|f(k)\| = \left[\sum_{k=a}^{b} |f(k)|^2\right]^{1/2} \tag{2}$$

The Norm of a vector in ℓ^2 space is defined by (2). A function belongs to the Hilbert Space ℓ^2 of discrete functions if the sum (2) is finite.

3) The Scalar Product in ℓ^2

$$(f,g) = \sum_{k=a}^{b} f(k)\overline{g(k)} \tag{3}$$

Scalar Product of two functions is defined by (3), where one of the functions is complex conjugate (marked by overline). This fact comes into effect if we deal with complex-valued sequences; however, Slepian Basis comprises real sequencies.

Aforementioned definitions are closely related to orthogonal and orthonormal bases. In an orthogonal basis $\{e_n(k)\}_{n=0}^{K-1}$ the every pair of non-identic sequences on interval (a,b) yields zero scalar product

$$(e_m, e_n) = 0 \quad \text{if} \quad m \neq n, \tag{4}$$

and the scalar products of identical sequences are non-zero

$$\|e_n(k)\| \neq 0. \tag{5}$$

We use the term orthonormal base if the basis is orthogonal and its elements $\{e_n(k)\}_{n=0}^{K-1}$ have unit norm $\|e_n\| = 1$. [2][3]

Discrete Fourier Basis, Discrete Cosine Basis, or Discrete Slepian Basis are all examples of orthogonal/orthonormal bases. Slepian Base is illustrated in the Figure 1.

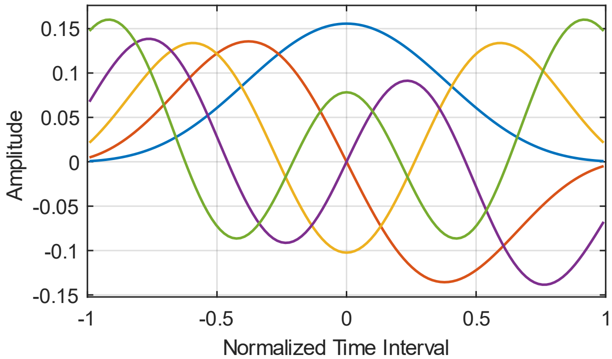


Figure 1: The Slepian Orthonormal Base for $K = 127$ samples. Sequences with orders from 0 to 4 are displayed.

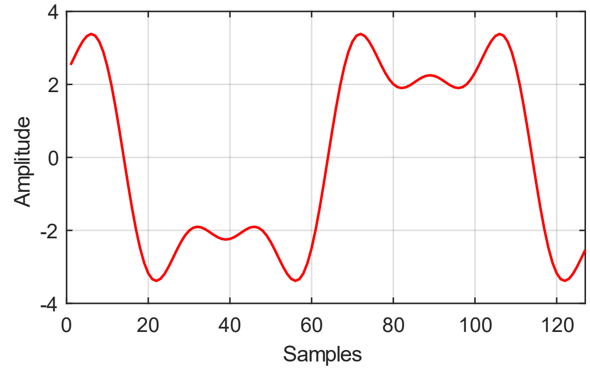


Figure 2: Waveform of the test signal (8).

B. Definition

Slepian Sequences are defined by the following convolutional equation

$$\sum_{l=1}^{K} \frac{\sin[2\pi W(k-l)]}{\pi(k-l)} \psi_n(K,W,l) = \\ = \lambda_n(K,W)\psi(K,W,k) \tag{6}$$

where k stands for the discrete time, K is the number of samples in the time domain, ψ_n is nth Slepian sequence, λ_n is its eigenvalue, and W is normalised bandwidth (a real number between 0 and ½). If we have normalized frequency domain, the value 1 corresponds to the sampling frequency, and in turn the maximal frequency admissible in signal spectrum is ½. [4][5]

The number of samples in time domain K is a parameter of the Slepian basis; it determines both the number and length of Slepian sequences. In this work is expected usage of the full Slepian Basis comprising K sequences, unless stated otherwise.

C. Spectrum

The measured signal $\{f(k)\}_{k=0}^{K-1}$ can be transformed into the Slepian Spectrum which is expressed by coefficients $\{a_n\}_{n=0}^{K-1}$. Considering a discrete-time sequence $f(k)$, we can obtain coefficients a_n from the second part of the double orthogonality of Slepian Sequences [4][6] as follows:

$$a_n = \sum_{k=1}^{K} f(k)\psi_n(k) = \lambda_n \sum_{k=-\infty}^{\infty} f(k)\psi_n(k) \tag{7}$$

Example of a Slepian Spectrum is made on a simple periodic signal

$$f(t) = 3\sin 2\pi t + \frac{3}{2}\sin 6\pi t + \frac{3}{4}\sin 10\pi t \tag{8}$$

which is illustrated in the Figure 2. The Slepian spectrum is illustrated in the Figure 3 and for comparison, there is also the Fourier Spectrum in the Figure 4.

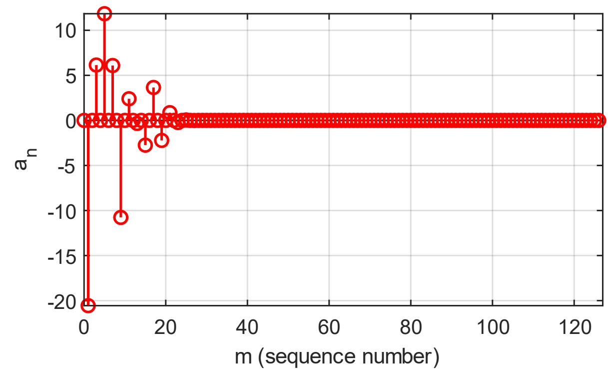


Figure 3: Slepian Spectrum of a harmonic the test signal (8).

The signal was deliberately sampled in such a way that one dicrete period of the signal does not contain an integer multiple of the fundamental frequency 2π rad/s. We can see that spectral coefficients are negligible for the 20th and higher orders of Slepain spectrum. That is also valid for the Fourier Spectrum, although Fourier Spectrum is symmetric and complex.

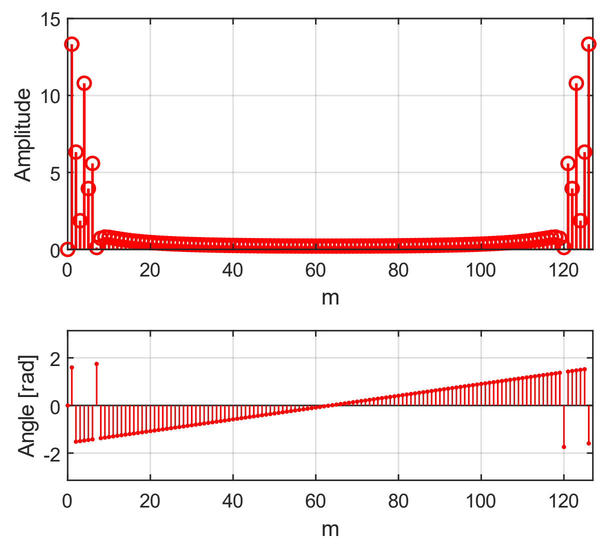


Figure 4: Fourier Spectrum of the periodic signal (8). Amplitude spectrum is in the upper chart, Phase spectrum is in the lower chart.

Considering the size of stored data in both spectra, the Fourier Spectrum requires only half of the coefficients stored in the computer's memory, because of its well-known symmetry when the processed signal is a real-valued sequence. In this case we need only 64 coefficients, because the sampled signal consists of 127 samples. However, the Fourier Spectrum is complex-valued and storing real and complex part of the number needs twice more memory. Indeed, storage of one half of the complex-valued Fourier Spectrum requires the same amount of memory as the full Slepian Spectrum.

III. The Slepian Transformation

Slepian Transformation (ST) is operation of computing spectral coefficients (7), Inverse Slepian Transformation (IST) reconstructs the signal from the spectrum. Process of IST is illustrated in the Figure 5, where approximation order corresponds with number of used products of Sequencies and coefficients a_n. The first approximation order represents IST of only the first spectral coefficient (and others equal to zero). As our first spectral coefficient is close to zero, the first approximation shows a constant sequence where amplitude approaches 0. The higher the approximation order, the more accurate the signal approximation is. Note that roughly from 20th-order approximation, there are almost no changes in the shape of signal, which is in conformity with Figure 3: coefficients a_n with order n higher than 20 are negligible.

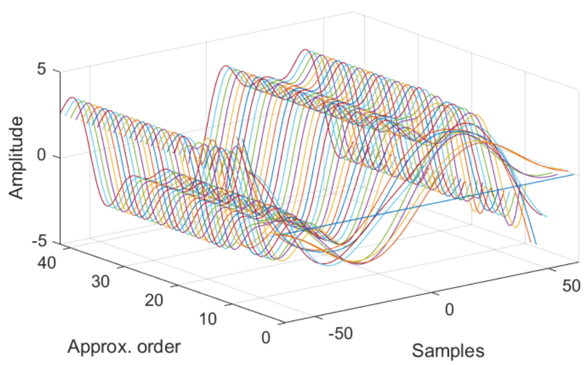


Figure 5 - Signal reconstruction using the Inverse Slepian Transform.

A. Compressed Spectrum – Inverse Transform Example

The hypothesis that high-order Slepian coefficients are negligible is tested by ST and IST. Example of signal employed for this task is depicted in Figure 6 and Figure 7 .

During the reconstruction, higher order coefficients are ignored, even though the full spectrum would have $K = 127$ coefficients. (The spectrum is computed from $K = 127$ signal samples). Parts of the signal are reconstructed with a truncated spectrum. Figure 8 shows the detail of signal approximation obtained from the truncated spectrum, using only ¼ of spectral coefficients. Truncated IST is compared with well-known and commonly used inverse Fast Fourier Transform (IFFT) and inverse Cosine Transform (ICT). For more information about CT see [7].

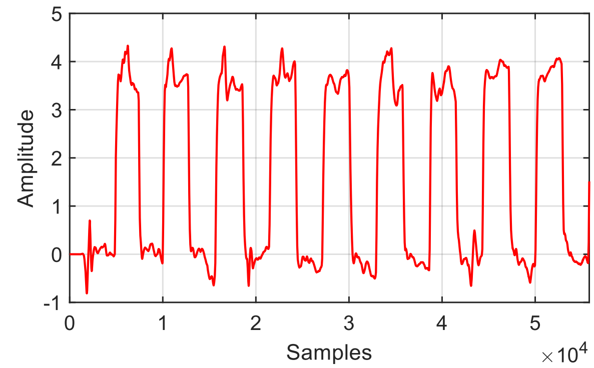


Figure 6: Example of a test signal (lateral vehicle position on a motorway).

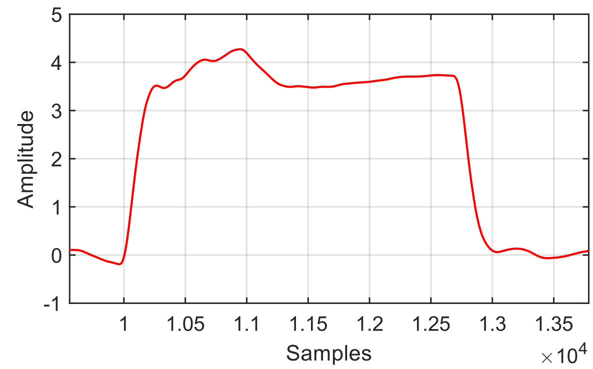


Figure 7: Example of a test signal (lateral vehicle position on a motorway), detail.

We can see that both the truncated IST and ICT can reconstruct the signal with high precision even if only a part of the spectrum is used. This might be useful to compress data in some special purposes.

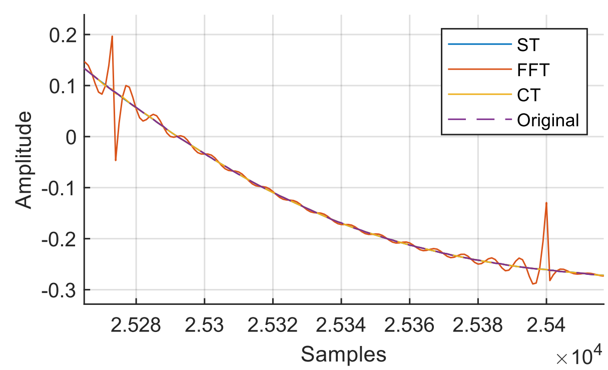


Figure 8: Detail of a compressed-spectrum signal reconstructions.

B. Compressed Spectrum – Inverse Transform Generalization

The process of IST with compressed truncated from Section III.A is generalized by the following algorithm:

1. Take the first spectral coefficient.
2. Execute the inverse transform (IST, ICT, or IFFT).
3. Compare it with the original signal, compute MSE (Mean Squared Error, [7]).
4. Add next spectral coefficient and go to 2.

5. If all coefficients from spectrum are used, go to 1 and continue with the next signal section of 127 samples.

6. If the whole signal is reconstructed, compute the relative global squared error

$$E_{\mathrm{R}} = \frac{\sum_{k=0}^{N-1} |f_m(k) - f(k)|^2}{\sum_{k=0}^{N-1} |f(k)|^2} \tag{9}$$

for each transform and each number of used spectral coefficients.

To compute the dataset of the following figure were used over 1,600 signal sections (from the signal in the Figure 6 or similar), each section being 127 samples long.

Figure 9 illustrates average relative global MSE of a Signal reconstruction in logarithmic axes of ST and FFT, CT to compare with. ST has the highest initial MSE because of the zero order sequency is not being constant (Slepian base has no constant sequency). In comparison, the FFT and CT both have the zero order sequency as a constant. However, after 10th coefficient used in the algorithm, ST is the most accurate with the lowest relative global MSE. The FFT MSE does not converge to zero as fast as other transforms which is caused by its characteristic “oscillation” on the edges of the interval (can be also seen in Figure 8).

It turned out that Fourier transform yields the slowest convergence of the signal approximation, as it is evident from the Figure 9. This corresponds with the Figure 8 where it is presented on a part of measured signal. On the other hand, ST yields the most accurate signal approximation even with only 10 first spectral coefficients used.

The FFT needs only ½ coefficients in comparison with ST or CT, but they are complex, as it was already discussed in Section II.C.

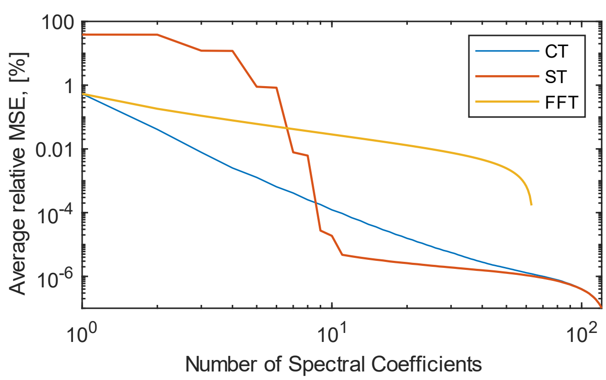


Figure 9: The relative global squared error (relative MSE) on used spectral coefficients in the signal approximation dependency.

IV. Conclusion

In this paper we briefly revisited the theory of orthogonal bases with the aim of finding bases most suitable for compression of data from a vehicle-driving simulator.

Three different signal decompositions were tested and compared, namely: FT, DCT, and ST. Our numerical experiments suggest that for the class of signals processed in this study Slepian transform affords the sparsest representation. This was illustrated by a number of mathematical tools: Plots of spectral coefficients corresponding to different bases, waveforms of signal approximations, and plots of relative MSE for increasing orders of signal approximation.

The authors are aware that there is room for further improvement in a number of ways. E.g., the free parameter of Slepian Basis (the product NW [5]) could be optimized for the given group of signals so as to yield sparsest possible representation when similar signals are processed.

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] C. T. Chen, “Signals and Systems”, 3rd Edition. New York: Oxford University Press, 2004.

[2] G. H. Hardy and W. W. Rogosinski, “Fourierovy řady”. Praha: STNL – Nakladatelství technické literatury, n. p., 1971.

[3] A. Kufner and J. Kadlec, “Fourierovy řady”. Praha: Academia, 1969.

[4] D. Slepian. “Prolate spheroidal wave functions, Fourier analysis, and uncertainty–V: the discrete case”. Bell System Technical Journal, vol.57, no.5, pp.1371-1430, May 1978

[5] “MathWorks Help Center: dpss”, The MathWorks, 2022. Available from: https://www.mathworks.com/help/signal/ref/dpss.html?searchHighlight=dpss&s_tid=srchtitle_dpss_1 (accsessed Dec. 26, 2022).

[6] O. Mihálik, D. Michalík, P. Jura and P. Fiedler, “Spectral Super-Resolution and Band-Limited Extrapolation Using Slepian Series”. In 16th IFAC Conference on Programmable Devices and Embedded Systems – PDeS 2019. IFAC-PapersOnLine (ELSEVIER). 52. IFAC, 2019. s. 388-393. ISSN: 2405-8963.

[7] K. R. Rao and P. Yip, “Discrete Cosine Transform: Algorithms, Advantages, Applications”. San Diego, CA: Academic Press, 1990.

Image demosaicing using Deep Image Prior

Peter Balušík
Dept. of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
230531@vut.cz

***Abstract*—The paper focuses on the problem of image demosaicing using the deep image prior. The deep image prior (DIP) is an uncommon concept that uses a generative neural network which, however, utilizes only the degraded image as the input for training. A novel method for image demosaicing is proposed, based on DIP, and it is compared with common demosaicing methods. In terms of the objective PSNR and SSIM values, the proposed method proved to be comparable with a widely used Malvar's demosaicing method. Nevertheless, subjectively, DIP produces demosaiced images comparable with the superior Menon's algorithm. Unfortunately, the proposed method turned out to be computationally immensely challenging.**

***Index Terms*—demosaicing, debayerization, color filter array, deep image prior**

I. Introduction

At present, most people own a mobile phone with a camera. Virtually everyone can say they have taken a photo at least once in their life. However, only few people understand how a digital color photo is acquired. To put it briefly, a photo comes into existence by light passing through the lens and projecting onto a flat image sensor with a *color filter*. After the image sensor has captured the different color components, *demosaicing* must begin. Demosaicing is a process that is usually built into the camera. It generates an image that can be viewed after taking a photo. If demosaicing was not present, the image would be obtained in the RAW format. A RAW image contains only the unprocessed information captured by the camera sensor.

The most common color filter used is the Bayer filter [1]. It is a color filter array (CFA) designed in such a way that each of the array cells allows solely the red, green, or blue (RGB) light component to pass. A usual configuration of such a filter can be seen in Fig. 1. It is made using a square pattern with two green elements, one blue and one red element. This pattern is then repeated to cover the whole sensor area. A commonly used arrangement is RG–GB (Fig. 1), but arrangements such as BG–GR, GR–BG, and GB–RG are also possible.

Naturally, a CFA can capture only one color value at each component. Demosaicing can be used to obtain full RGB values of each component. Therefore, demosaicing can be understood as a process of estimating the missing color values of a CFA. In other words, it is a special type of color interpolation. Many different demosaicing methods exist. One of the simplest and quickest methods is the bilinear interpolation. Other, more effective methods include more complex techniques such as Malvar's [2] and Menon's demosaicing methods [3].

This work was supported project No 23-07294S of the Czech Science Foundation (GAČR).

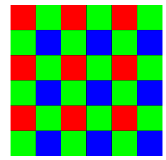


Fig. 1. Schema of Bayer filter with RG–GB arrangement.

Malvar's method uses the bilinear interpolation but corrects its result using the value of the gradient multiplied by a gain factor. The gradient is calculated using predefined filters which are described in [2]. The gradient correction results in sharper images. Additionally, it slightly reduces aliasing artifacts originating due to using the bilinear interpolation.

Menon's approach slightly differs from those mentioned above. It uses a directional interpolation of the green pixels along both horizontal and vertical directions and creates two corresponding "green" images. Following the interpolation, a choice is made between the two green images using two classifiers [3]. Subsequently, the red and blue components are interpolated. Besides information from the Bayer filter, the obtained green image and the two classifiers help with the reconstruction of these components. This results in more accurately reconstructed color channels with significantly less pronounced artifacts.

Deep neural networks (DNN) have also been exploited in demosaicing. Klatzer et al. [4] created a DNN that was able to learn its weights on the training image dataset such that the network performed joint demosaicing and denoising. Kurniawan et al. [5] utilizes the deep image prior (DIP) concept [6] as the present paper does, but it does so for random CFAs containing white elements in addition to RGB.

This paper focuses on an approach to demosaicing related to [5], but using the DIP idea in a much more straightforward way. We adopt the DIP network to solve the image demosaicing problem for its use with standard RGB CFA arrays, and perform comparisons of the restoration quality against the other standard methods.

II. Deep Image Prior

The Deep image prior [6] is a concept used for the reconstruction of degraded images. It restores the images using a generative convolutional neural network (CNN). The main difference from common reconstructive algorithms is that

DOI 10.13164/eeict.2023.17

DIP uses a network that is completely *untrained*. Given the structure of the network, its only input information is an image damaged or degraded in some way. The task of the network is to recover the image as well as possible. All of the weights of the network are randomly initialized. The weights of the network, serving as a parameterization of an image, are trained such that the network learns a conversion of a completely random initial image into the degraded image provided. The size of the network is usually great enough to allow such an unusual transformation after a sufficient number of learning epochs. The key observation of [6] is that along that learning path, a restored image appears. It is only necessary to *stop learning* at the right point. This way, DIP proves that the architecture of the neural network fundamentally influences the results (and needs to be taken into consideration in the future of deep learning). The DIP applied to image processing tasks typically utilizes the U-net-like "hourglass" architecture, also called encoder–decoder with skip connections.

Several applications of this neural network are described in [6]. These applications include denoising, super-resolution, inpainting, image enhancement, removal of JPEG-compressed artifacts. In some applications, the visual quality of the restored images is comparable with the state-of-the-art methods, which are *trained* in most cases. The downside of this concept is that the optimization process is very slow, compared with the other networks which—once learned—are fast.

A deep neural network used for image generation can be considered a parametric function $x = f_\theta(z)$ that maps a code vector z to an image x. In the case of DIP, the code vector z is randomly initialized but fixed in the course of the learning epochs. The parameter θ incorporates all the weights and biases present in the network.

Reconstruction tasks can be formulated as energy minimization problems:

$$x^* = \min_x E(x; x_0) + R(x), \tag{1}$$

where $E(x; x_0)$ is the data term that is chosen according to the required application, x_0 represents the damaged image and $R(x)$ is an explicit regularizer. The regularizer is not tied to particular applications; commonly it characterizes naturally looking images. But in the case of DIP, R is replaced by an implicit prior captured by the neural network itself:

$$\theta^* = \operatorname*{argmin}_\theta E(f_\theta(z); x_0), \tag{2}$$

where the minimizer θ^* can be found using standard optimization techniques. Once θ^* is obtained, the final recovered image can be reconstructed by the application of the network to the initial code, i.e. $x^* = f_{\theta^*}(z)$. As already mentioned, the ability of the network to (over)fit and thus result in the corrupted image itself (i.e. $x_0 = x^*$) requires stopping the training process after a proper number of epochs [6].

III. Method

As mentioned above, one of the applications of DIP in [6] is inpainting. Inpainting is a process where damaged or missing portions of a digital image are repaired (repainted). The authors of [6] propose the following approach to solve this problem. Let x_0 be the depleted image of size $H \times W$, and $m \in \{0, 1\}^{H\times W}$ the binary mask indicating missing pixels. Then the image can be restored using the following data term (after parameterization of the network):

$$E(f_\theta(z); x_0) = \|(f_\theta(z) - x_0) \odot m\|^2, \tag{3}$$

where $\odot$ is the Hadamard product. The inclusion of the data prior is crucial, given that the energy (data) term alone would not be sufficient to spread information about the color values to the missing sections. If the data prior was not present, the resulting image would simply not change after optimization over pixel values x. Again, DIP uses an implicit data prior that is given by the optimization of (3) with respect to parameters θ.

In this paper, an approach similar to inpainting is proposed for demosaicing. Instead of the binary mask that takes on the values 0 or 1, a Bayer mask involving RGB channels is used. This paper works with an RG–GB Bayer mask but other combinations would result only in the change of the mask. If the binary mask m is replaced by a Bayer mask M_{CFA}, the data term for demosaicing becomes

$$E(f_\theta(z); x_0) = \|(f_\theta(z) - x_0) \odot M_{\text{CFA}}\|^2. \tag{4}$$

Again, the data prior is present through the optimization of a data term similar to (3).

IV. Experiments and Results

A. Data set

All 24 images from the traditional Kodak image data set [7] were used in the experiments. The size of these images was 768×512 or 512×768, depending on the image orientation. The images come in the uncompressed PNG-24 file format with 8 bits per channel.

B. Artificial mosaicing

Real "RAW" images, in practice acquired by the camera sensor, had to be simulated. Therefore, an artificial bayerization, or mosaicing, was performed on the Kodak images. An RG–GB Bayer mask was used to create R, G, B undersampled images.

C. Parameters of the neural network

An encoder–decoder architecture similar to the U-net [8] with five downsampling and five upsampling convolutional layers was chosen for the demosaicing method. Even though skip connections led to a faster optimization, they also brought undesirable artifacts, at least with the parameter settings used in the experiments. The learning rate used was 0.001 and the network weights were randomly initialized with Gaussian noise. The counts of feature maps in both the upsampling and downsampling convolution layers were as follows: 64, 64, 128, 256, and 512. Compared with the inpainting experiment presented in [6], the number of feature maps was the most distinctive setting of the proposed network. The feature maps were obtained by a 3×3 convolution with zero padding, followed by a LeakyReLU activation function. The Adam optimizer [9] was used as the optimization algorithm. The number of epochs was chosen to be 2000 since the highest PSNR values most frequently appeared around 2000 epochs. Also, the maximum SSIM values were observed close to 2000 epochs.

D. Evaluation

The proposed demosaicing method was compared with the bilinear interpolation, to Malvar's and Menon's methods, and to the `demosaic` function in Matlab (very similar to Malvar's demosaicing method). As the objective scores for the comparison, the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM) [10] were used. It is worth noting that the higher the PSNR value, the better. The maximal SSIM value is 1. The PSNR and SSIM values were obtained using the `scikit-image` [11] Python package.

It is worth noting that the classic demosaicing methods mentioned above are purely deterministic, while randomness of several kinds affects the output of the DIP-based method: First, the code vector z is newly generated each time the optimization is run. Second, the network is initialized with random weights θ, meaning that every time the program runs, a different local minimum of the loss function can be found. Finally, the optimization algorithm utilizes the stochastic gradient descent [9], which may also lead to a perturbation in the final solution. As a consequence, the proposed method is not perfectly stable in terms of the PSNR/SSIM over multiple runs, since the results vary for each run of the program. However, the variation is not severe: Across the test images, the PSNR values of individual program runs differ by ± 0.3 on average. As for the SSIM values, the variation is negligible (± 0.002).

E. Findings

1) Comparison of PSNR and SSIM values: The proposed method proved to be on a par with the `demosaic` function in Matlab and close to Malvar's democaicing method, with an average PSNR of 33.7 dB and an average SSIM of 0.967 (see Table I). Furthermore, is was marginally better than bilinear interpolation. Menon's method proved to be the superior method among all methods. A comparison of PSNR and SSIM values can be seen in Fig. 2.

2) Subjective visual comparison: Visually, the proposed method performs significantly better than any other method except for Menon's demosaicing. Most images demosaiced by the proposed method are indistinguishable from those produced by Menon's method. A visual difference reveals itself rarely, for example, in the case of the image no. 19, which is shown in Fig. 3. No formal subjective test was performed to prove these impressions statistically, however.

3) Additional improvements: The value of the objective criteria typically slightly oscillates in the course of training epochs. Thus, the proposed method can be improved by averaging the images from the last few epochs. This improvement comes basically with no additional cost. In terms of the number of images averaged, 50 images proved to be a sensible choice with regard to objective quality. Averaging more than 50 images did not bring any significant improvement, doing so only (very slightly) increases the time of execution. The described development results in a PSNR improvement of 0.4 dB on average. The SSIM improves slightly too; however, the improvement is negligible (0.0019).

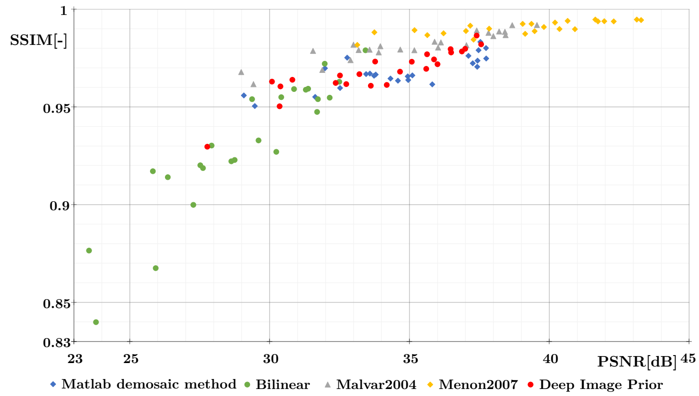


Fig. 2. A scatter plot of the PSNR and SSIM values of all images and methods. The vertical axis represents the SSIM values while the PSNR values are displayed horizontally.

TABLE I
TOP BLOCK: AVERAGED RESULTS OF TRADITIONAL METHODS. BOTTOM BLOCK: AVERAGED RESULTS OF 10 RUNS OF THE PROPOSED ALGORITHM ON THE IMAGES FROM THE KODAK DATA SET.

	PSNR [dB]	SSIM [-]
Bilinear interpolation	29.1635	0.93098
Malvar's method	35.1601	0.98111
Menon's method	**39.0132**	**0.99049**
Matlab `demosaic` function	34.6454	0.96777
Value at 2000 epochs	33.7053	0.96685
Oracle (without averaging)	34.0019	0.96821
Oracle (with averaging)	34.1320	0.96875
Images 1999 & 2000 averaged	34.0039	0.96797
Images 1900 & 1901 averaged	34.0095	0.96780
Best run (without averaging)	33.9096	0.96845
Best run (with averaging)	**34.3156**	**0.97022**
Images 1991–2000 averaged	34.0862	0.96841
Images 1981–2000 averaged	34.0987	0.96849
Images 1971–2000 averaged	34.1064	0.96856
Images 1961–2000 averaged	34.1133	0.96861
Images 1951–2000 averaged	34.1177	0.96863
Images 1901–2000 averaged	34.1226	0.96861

4) Downsides: The crucial downside of the proposed method and the Deep Image Prior as a whole is that it is extremely computationally demanding compared with other methods. Averaged over all test images, the execution time of classic demosaicing methods is below one second, with the fastest being the bilinear interpolation at around 0.4 seconds. The proposed method takes, around 90 seconds on average. Unlike with the other methods, the time is heavily dependent on the computer hardware. The shortest reconstruction times are obtained when a GPU is utilized. Our results were acquired using the NVIDIA Tesla V100S PCIe 32 GB graphics card. It is important to consider that the times presented hold for

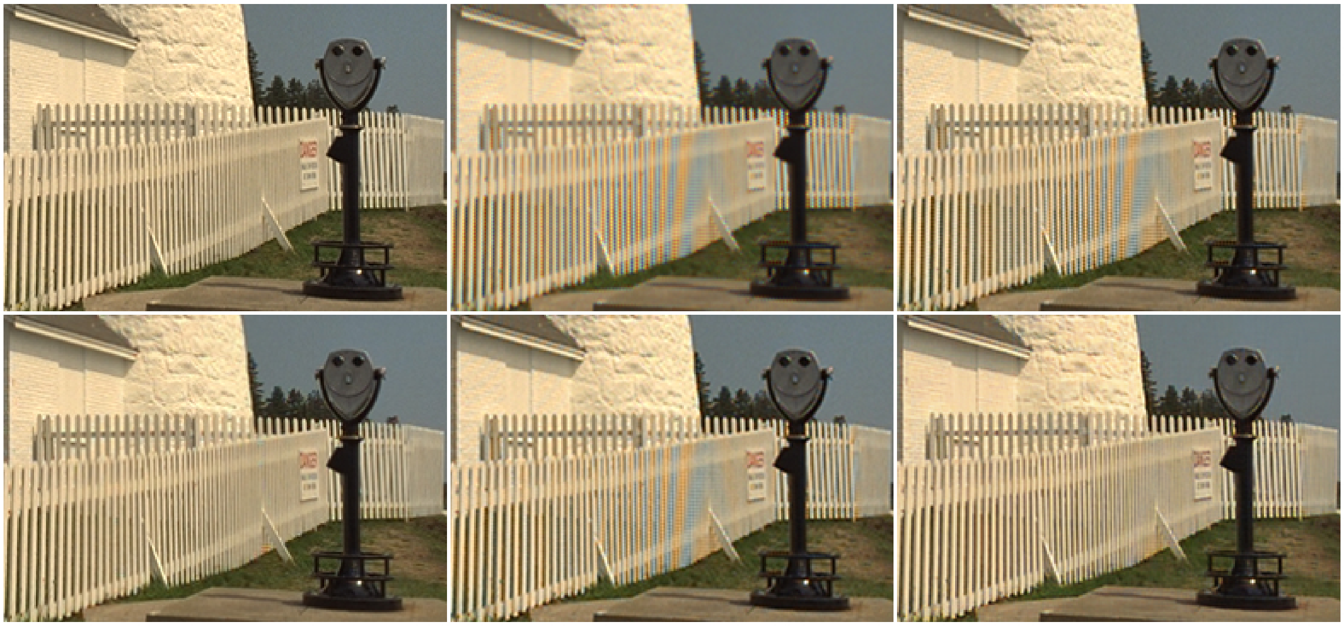


Fig. 3. Visual comparison of all methods on the fence from the Kodak image no. 19 – the colored stripes on the images represent color artifacts mentioned in the text. a) the ground truth; b) bilinear interpolation; c) Malvar's method; d) Menon's method; e) Matlab `demosaic`; f) the proposed method.

768×512 images, which are relatively small in the context of today's cameras. Larger images would require even more computation time.

5) Surprises: To show the maximal capabilities of the proposed method, an oracle approach was selected. In this case, the oracle approach means that the data is treated as if the ground truth image was known. As a result, the images with the maximum PSNR/SSIM value could be selected; in our case, the best image (in terms of PSNR/SSIM) from epochs 1900–2000 was chosen. The greatest surprise came with averaging the last few images. Averaging only the last two images proved to be on a par with the oracle approach, at least in terms of PSNR. Taking it even further, averaging the last 10 images came with better results than the oracle approach. The comparison of PSNR and SSIM values of different numbers of averaged images and other noteworthy statistics can be seen in Table I.

V. Conclusion

The paper proposed a demosaicing method based on the DIP concept. In terms of PSNR and SSIM values, it delivered results very similar to those of the classic Malvar's method. Visually, DIP did not surpass only Menon's method, whose results are comparable with DIP. The most surprising fact was that averaging the last few images of the optimization process brought very solid results, even when compared with the oracle approach. The resulting images from the Kodak data set are available at `https://github.com/sedemto/results_dip`. It is worth mentioning that even though DIP has shown very good visual results, it cannot be used in practice due to the fact that it is very computationally demanding.

Future development could involve the inclusion of an additional regularization of the network to tackle aliasing artifacts even more strongly. Also, the joint processing of the RGB channels by the DNN could be revised. Finally, in difference to most of other methods, the DIP-based demosaicing could be straightforwardly adapted to the case of a random CFA, where aliasing issues are automatically restrained.

Acknowledgment

This work was supervised by dr. Pavel Rajmic, Signal Processing Laboratory, Brno University of Technology.

References

[1] B. E. Bayer, "Color imaging array," US3971065A, Jul. 20, 1976.

[2] H. S. Malvar, L. He, and R. Cutler, "High-quality linear interpolation for demosaicing of Bayer-patterned color images," in *2004 IEEE Intl. Conference on Acoustics, Speech, and Signal Proc.*, May 2004, vol. 3.

[3] D. Menon, S. Andriani, and G. Calvagno, "Demosaicing With Directional Filtering and a posteriori Decision," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 132–141, Jan. 2007.

[4] T. Klatzer et al., "Learning joint demosaicing and denoising based on sequential energy minimization," IEEE *Intl. Conference on Computational Photography* (ICCP), USA, 2016

[5] E. Kurniawan, Y. Park, and S. Lee, "Noise-Resistant Demosaicing with Deep Image Prior Network and Random RGBW Color Filter Array," *Sensors* 22, no. 5: 1767, 2022

[6] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep Image Prior," *Intl. Journal of Computer Vision*, vol. 128, no. 7, pp. 1867–1888, Jul. 2020.

[7] M. Bakr, "Kodak Lossless True Color Image Suite." Oct. 14, 2022. Accessed: Mar. 08, 2023. [Online].

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention 2015*, Oct. 2015, pp. 234–241.

[9] D. P. Kingma and L. J. Ba, "Adam: A Method for Stochastic Optimization," in *Intl. Conference on Learning Representations (ICLR)*, 2015.

[10] Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Proc.*, vol. 13, no. 4, 2004.

[11] S. van der Walt et al., "scikit-image: Image processing in Python," *PeerJ*, vol. 2, p. e453, Jun. 2014.

Compatibility of piezoelectric semicrystalline polymer and osteoblastic cells

Tereza Havlíková
Department of Physics, Faculty of Electrical Engineering and Communication
Brno University of Technology
Technická 2848/8, 61600 Brno, Czech Republic
Tereza.Havlikova1@vut.cz

***Abstract*—For the improvement of tissue regeneration, functional polymer scaffolds that can mimic the extracellular matrix of human tissue are extremely desirable. Depending on the tissue that the scaffolds are intended to resemble, these scaffolds must satisfy highly precise requirements such as non-toxicity or fibrous structure. High piezoelectricity and hydrophilicity also proved to have convenient effects. Due to their innate capability to create surface charges under mechanical stress, piezoelectric materials such as polyvinylidene fluoride (PVDF) become excellent candidates for creating functional scaffolds. It is desired for PVDF to also have hydrophilic properties. Otherwise, it could prevent adequate cell adhesion and growth necessary for the construction of biomimetic scaffolds. For this study, electrospinned PVDF nanofibers covered by human or mouse osteoblasts were subjected to Raman spectroscopy, measurement of the contact angle of the liquid wettability on the sample surface to observe hydrophobicity and hydrophilicity, and scanning electron microscopy (SEM) to assess properties of the material in relation to oxygen plasma treatment.**

Keywords—scaffold, PVDF, osteoblasts, piezoelectricity, hydrophilicity, hydrophobicity, adhesion, wettability

I. Introduction

The potential use of artificial scaffolds in regenerative medicine is rising to the center of attention as a result of the ongoing quest for alternative therapies. It appears that biomimetic tissue replacements have the capacity to solve the continuing crisis of tissue or organ shortage. The scaffolds are used for simulating the extracellular matrix that cells make for their support. It is a place where the cells grow, proliferate, and interact with each other.

Due to their processability, morphological traits, and wide range of potential applications in the field of electronic devices and biomedical science, advanced synthetic polymeric biomaterials have gained attention in recent years. Among them, PVDF with a natural piezoelectric effect has a variety of qualities that make it a flexible and functional biomaterial for cell attachment and subsequent tissue development [3].

Certain criteria for the surface features and microarchitecture of the biomaterial must be considered depending on the targeted application. Thus, as they will affect the way in which cells and surfaces of the scaffolds interact, the physical, chemical, and hydrophilic or hydrophobic properties of the scaffold must be well recognized [1].

The overall layout of nanofibers is crucial for creating a solid scaffold for cells to grow on. It is well known that some cell types depend on the right alignment of nanofibers. With electrospinning, the ideal arrangement and fiber structure can be produced by controlling the parameters of the working process such as the speed of the collector, voltage, and flow rate of the solution [2]. The electrospinning technology is also superior in terms of versatility of usage, efficiency, and consistency in the quality of the outcoming materials [3].

Electrospinning is also one of the most effective methods for synthesizing piezoelectric materials at the nanoscale. Piezoelectric polarization in polymer materials is induced by mechanical stress and is crucial for specific material implementations such as the formation of functional scaffolds for cells to grow on. The PDVF has three distinct crystalline phases - α (alpha), β (beta), and γ (gamma), with different chain conformations. The phase that forms in the majority of polymers is the α-phase which is electrically inactive. However, the polar β-phase is essential for the PVDF material because it exhibits the most potent piezoelectric property. The pyroelectric, piezoelectric, and ferroelectric capabilities of the β-phase PVDF as well as its inherent low weight, mechanical flexibility, and ease of manufacturing led to extensive research into these materials. Electrospinning is suitable for the production of PVDF material with a high β-phase [2,3].

Another important factor in choosing the most appropriate material is hydrophobicity or hydrophilicity. It is a property of surfaces and their molecules that is manifested by the reluctance to interact with water. Highly hydrophobic materials hinder cell adhesion and the formation of functioning tissue. Therefore, it is desired to alter the chemical surface of the material by methods such as oxygen plasma treatment to render the material hydrophilic [1].

II. Materials and Methods

A. Materials and their Parameters

The nanofibers used in this study were electrospinned by a 4SPIN machine (Contipro, Dolní Dobrouč, Czech Republic) with a cylindrical collector covered with aluminum. The polymer solution passed through the needle emitter at a flow rate of 30 μl per minute. While the collector spun at a speed of 2000 or 300 rounds per minute, the polymer solution was drawn to the surface of the collector and formed a thin layer of nanofibers. The voltage between the emitter and the collector was 50 kV.

The participating material was a 20% solution of PVDF with a molecular weight of 275,000 g/mol. The solution was dissolved at a rate of 7:3 in a solvent of dimethyl sulfoxide and acetone. The final mixture was heated for 24 hours on a stirrer at a temperature of 80 °C and a speed of 200 rpm.

One-half of the used nanofiber samples were treated with oxygen plasma by using the Piezobrush PZ3 device (Relyon Plasma GmbH, Regensburg, Germany) as means of reducing the hydrophilicity of the surface of the material.

Coater EM ACE600 instrument (Leica Microsystems, Germany) was used for coating the surface with a 15 nm thick layer of carbon. The coating prevents the electric charge from accumulating on the material´s surface add fixates the fibers in place during scanning electron microscopy. Without the coating, it would be difficult to focus on the nanofibers and scan them at high magnification.

B. *Cultivation of the Cells*

Cells used for this work were human Saos-2 derived from osteosarcoma and mouse osteoblastic MC3T3 cells. The changes in the standardized process of cell cultivation were made based on prior research using osteoblasts [5].

A Dulbecco's Modified Eagle's medium supplemented with 10% fetal bovine serum and 5% penicillin/streptomycin (50 UI mL^{-1} and 50 g mL^{-1}) was used to maintain human osteosarcoma Saos-2 cells (ATCC® HTB-85TM) at 37 °C in a humidified 5% CO2 incubator. Trypsinization using a 0.25% trypsin-EDTA solution was used to extract the cells. Saos-2 cells were seeded at the density of 1 × 103 · mL^{-1} onto the sterile PVDF samples (1 × 1 cm).

Mouse MC3T3 osteoblasts were maintained in a culture flask with Alpha MEM Eagle culture medium. After the culture process was finished, the cells were separated into test tubes. Cells were then plated on PVDF samples (1 × 1 cm), placed in a Petri dish, and covered with Alpha MEM Eagle again. Petri dishes were placed into an incubator at 37 °C and 5% CO2 until the cells were attached to the material.

C. *Methods Used for Assessing Properties of the Samples*

1) Raman Spectroscopy: The three crystalline phases of the studied material were determined from the Raman spectrum using the WITec Alpha 300R device (WITec, Ulm, Germany). During the spectroscopy 532 nm green laser was used with a power of 6 mW. The number of accumulations was 10x at an integration time of 10 s using a lens with a magnification of 50x. The samples used for the spectroscopy were 2 sets of nanofibers that were electrospinned with different values of the collector´s spinning velocity. The first set of nanofibers was spun at 300 rpm and the second was spun at 2000 rpm.

2) Hydrophobicity and Hydrophilicity: A See System E (Advex Instruments, Brno, Czech Republic) device was used for the measurement, and See Software 7.0 was used to assess the contact angle from the photographs. Using a dosing micropipette, ten drops of distilled water at 3 µm size were progressively added to the surface of the studied nanofibers. At time t = 4 s, the contact angle was measured from a photograph by producing a circle from three chosen spots accurately replicating the drop form. For the observation of wettability were chosen 2 sets of nanofibers with different modifications in the case of plasma treatment.

3) Scanning Electron Microscopy (SEM): The fiber structure of the polymer material and interaction between the nanofibers and the osteoblastic cells were observed by scanning electron microscopy on a LYRA3 microscope (Tescan, Brno, Czech Republic). For all the observations following working parameters were set: detector SE, voltage 5kV, working distance 9 mm, and magnification from 3 kx to 185 kx.

III. Results

A. *Investigation of PVDF´s Crystalline Phases*

Raman spectroscopy is based on how light interacts with chemical bonds in a substance. By assessing peak positions and relative peak intensities it is possible to uncover the chemical structure of the material, crystalline phases, intrinsic stress or contamination, and impurity [3].

Both samples subjected to the Raman spectroscopy exhibit presence of the β-phase. The most characteristic peaks for the β-phase are at 840 cm^{-1} and 1275 cm^{-1} [2]. The first sample in Fig. 1a displays Raman spectrum of undirected nanofibers electrospinned at 300 rpm and presents β-phase peaks at 839 cm^{-1}, 882 cm^{-1}, and 1430 cm^{-1}. The second sample in Fig. 1b with directed nanofibers electrospinned at 2000 rpm exhibits the same peaks but in greater proportion and displays in addition 2 more β-phase peaks at 1074 cm^{-1}, and 1274 cm^{-1}.

The measurement results suggest that the material with directed nanofibers displays a higher β-phase than the material with undirected nanofibers which is the most important phase due to its piezoelectricity.

B. *Hydrophilic and Hydrophobic Properties of PVDF*

The water contact angle, which is the angle between the tangent to the liquid-vapor interface and the solid surface at the three-phase contact line, is a typical way to describe wettability. The contact angle lowers as the surface energy of the tested surface increases since there is a direct connection between the two. The hydrophobicity interval for the water contact angle is set from 90° to 150°. The values belonging to the hydrophilicity interval range from 0° to 90° [2].

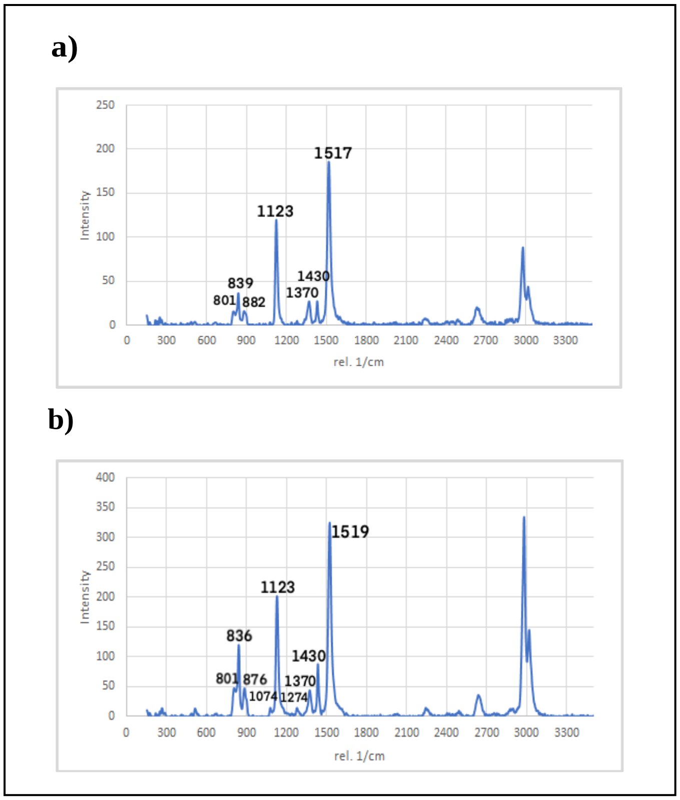


Fig. 1.: Raman spectroscopy: a) undirected PVDF nanofibers electrospinned at 300 rpm, b) directed PVDF nanofibers electrospinned at 2000 rpm.

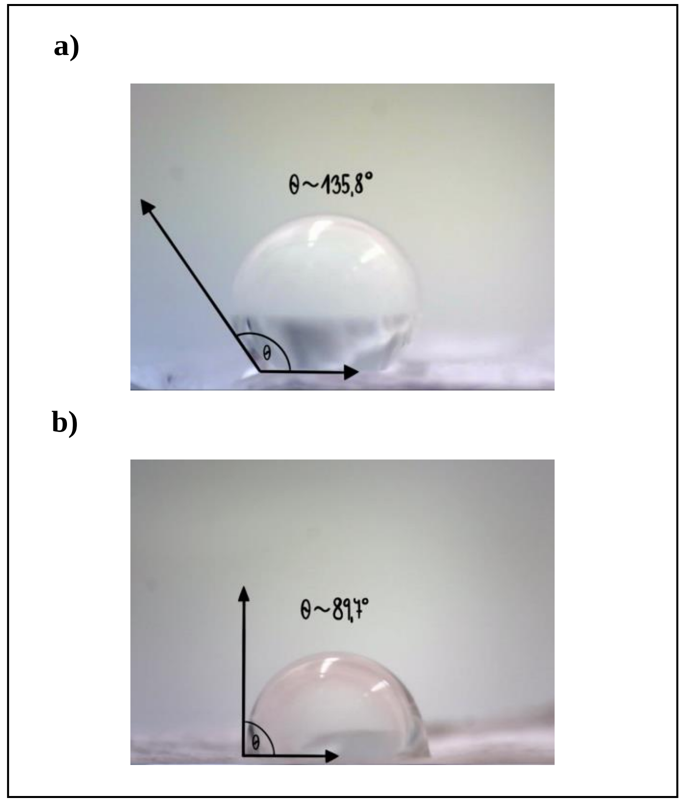


Fig. 2.: Measurement of water contact angle: a) sample of PVDF nanofibers non-modified by plasma, b) sample of PVDF nanofibers modified by plasma.

The surface water contact angle of the PVDF nanofibers non-modified by plasma in Fig. 2a has an average value of 125,6° acquired from 10 measurements at t = 4s. After t > 600s the material did not absorb more of the liquid. Based on the data, the non-modified PVDF material may be classified as hydrophobic. However, the surface water contact angle of the nanofibers modified by plasma in Fig. 2b has an average value of 107,8° with the best result at 89,7° which belongs to the hydrophilic interval. After t > 600s part of the droplet´s volume was adsorbed into the material.

C. Cell Adhesion Concerning Plasma Treatment of the PVDF Nanofiber Scaffold

Oxygen plasma treatment was used to alter the wettability of PVDF nanofiber scaffolds' surfaces. It was established that this method had no discernible effect on fiber size or electroactive β-phase content. Plasma treatment primarily results in surface chemistry changes, such as the addition of oxygen and the release of fluorine atoms, which considerably alter the wettability of polymer membranes by lowering their total surface tension. Due to the integration of oxygen and hydrophilic functional groups, the use of oxygen plasma on various polymer materials has shown encouraging results for stimulating cell development in the fabricated scaffold [4]

Scanning electron microscopy revealed a significant change in the roughness of the individual nanofibers in Fig. 3a displaying material modified by plasma in comparison to Fig. 3b showing non-modified material. However, the high energy supplied by the plasma source in some cases led to melting and joining of the neighboring fibers. Regardless of the fiber jointing, the modified scaffold remained much of the original open spaces between the fibers. The modified sample´s new characteristics in Fig. 3c in comparison to the non-modified sample in Fig. 3d enhanced the integration of the cells into the scaffold, their individual adhesion to the fibers, and the overall formation of cell clusters.

Observations done by SEM offered useful insight into the complicated relations between the surface morphology of PVDF scaffolds, their physical properties, and the mechanical support they provide to the growing osteoblastic cells. The human osteoblasts integrated into the material and attached well to the individual fibers even though the material was not modified by plasma, but it appears that the mouse osteoblasts used the potential of the scaffold even better due to its modification.

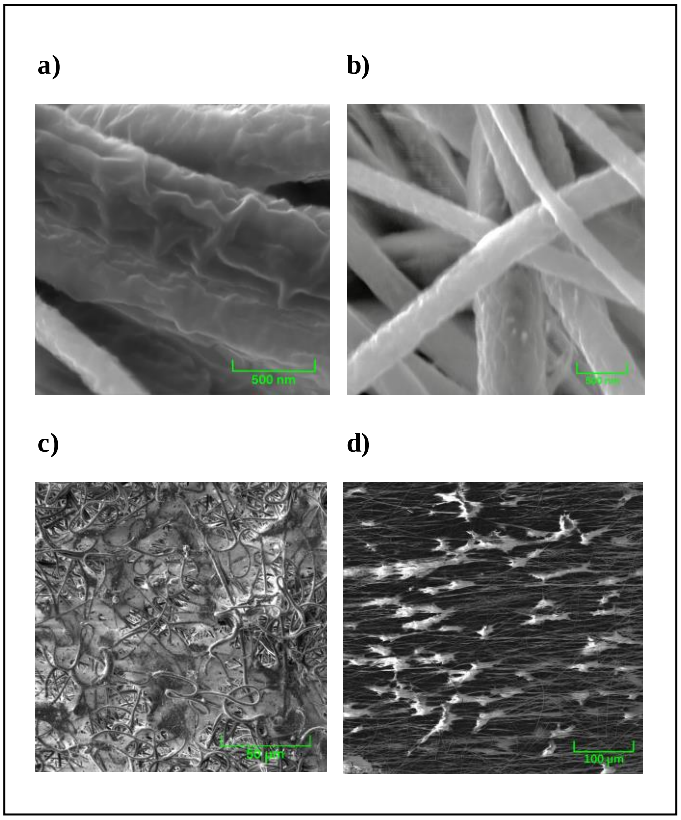


Fig. 3.: Scanning electron microscopy: a) surface of PVDF nanofibers modified by plasma, b) surface of PVDF nanofibers non-modified by plasma, c) osteoblasts on PVDF nanofibers modified by plasma, d) osteoblasts on PVDF nanofibers non-modified by plasma.

IV. Conclusion

During the process of preparation, PVDF nanofiber scaffolds and cell samples were successfully made. The planned measurements of the physical and morphological characteristics of the PVDF material were effectively executed and brought interesting results.

Raman spectroscopy proved the existence of crystalline phases in PVDF nanomaterial made in process of electrospinning and displayed a difference between nanofibers made with a distinct speed of the collector. The directed nanofibers were made at a speed of 2000 rpm and displayed greater relative peak intensities in β-phase than the undirected nanofibers made at 300 rpm.

During the measurement of the contact angle of the liquid wettability on the sample surface to observe hydrophobicity and hydrophilicity was non-treated PVDF assessed as hydrophobic. However, oxygen plasma treatment proved to be useful as means of lowering the liquid contact angle and rendering the PVDF hydrophilic. For further strengthening of the material´s hydrophilicity, it would be needed for an increase in used plasma power. This increase would, unfortunately, cause larger melting and joining of the nanofibers. This phenomenon could hinder cells´ integration into the scaffold because the melted nanofibers would not leave enough space for the cells to grow between them.

Minor lowering of the material´s hydrophobicity proved to have a positive effect on cell adhesion and integration. Yet the cells also integrated into the non-treated material with high hydrophobicity. This leaves a question if it is necessary to lower the liquid contact angle in every case of designing a suitable scaffold for tissue engineering or if it depends on the cell type or type of tissue the scaffold is designed to resemble. Despite that, the overall best results are reached with a PVDF nanomaterial electrospinned at high collector speed, with strong β-phase and treated with oxygen plasma for a minor lowering of the material´s hydrophobicity.

This work served as fundamental research on osteoblasts´ compatibility with a polymeric scaffold. Follow-up research could investigate further the need for PVDF to have hydrophilic properties and the reaction of the scaffold and the growing cell tissue to exposure to an electric field.

Acknowledgment

I want to express my gratitude to Ing. Nikola Papež, Ph.D., my adviser, for his insightful suggestions. I would also like to thank doc. Mgr. Zdenka Fohlerová, Ph.D., and Ing. Pavel Kaspar, Ph.D. for guidance with the sample preparation.

References

[1] Dumitrescu, L.N., Neacsu, P., Necula, M.G., Bonciu, A., Marascu, V., Cimpean, A., Moldovan, A., Rotaru, A., Dinca, V., Dinescu, M., “Induced, Hydrophilicity and In Vitro Preliminary Osteoblast Response of Polyvinylidene Fluoride (PVDF) Coatings Obtained via MAPLE Deposition and Subsequent Thermal Treatment,” Molecules 25, no. 3: 582., 2020, https://doi.org/10.3390/molecules25030582

[2] P. Černohorský, T. Pisarenko, N. Papež, D. Sobola, S. Ţălu, K.Částková, J. Kaštyl, R. Macků, P. Škarvada, and P.Sedlák, "Structure Tuning and Electrical Properties of Mixed PVDF and Nylon Nanofibers," Materials 14, no. 20: 6096, 2021 https://doi.org/10.3390/ma14206096

[3] Ruan, Liu Xia, Xiannian Yao, Yufang Chang, Lianqun Zhou, Gaowu Qin, and Xianmin Zhang, "Properties and Applications of the β Phase Poly(vinylidene fluoride)," Polymers 10, no. 3: 228, 2018 https://doi.org/10.3390/polym10030228

[4] D.M. Correia, C. Ribeiro, V. Sencadas, G. Botelho, S.A.C. Carabineiro, J.L. Gomes Ribelles, S. Lanceros-Méndez, “Influence of oxygen plasma treatment parameters on poly(vinylidene fluoride) electrospun fiber mats wettability,” Progress in Organic Coatings 85, 2015, https://doi.org/10.1016/j.porgcoat.2015.03.019

[5] L. Hernández-Tapia, Z. Fohlerová, J. Žídek, M. A. Alvarez-Perez, L. Čelko, J. Kaiser, and E. B. Montufar, "Effects of Cryopreservation on Cell Metabolic Activity and Function of Biofabricated Structures Laden with Osteoblasts," Materials 13, no. 8: 1966, 2020, https://doi.org/10.3390/ma13081966


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

Analog-digital module for electrical impedance tomography

1st David Zimniok
Department of Radio Electronics
FEEC, Brno University of Technology
Brno, Czech Republic
xzimni00@vut.cz

2nd Jan Mikulka
Department of Theoretical and Experimental Electrical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
mikulka@vut.cz

***Abstract*—The present text introduces the concept of data collecting (voltages and currents) from a phantom connected to electrical impedance tomography. The concept involves a method for signal multiplexing from up to 256 electrodes and the current source for such probes. Considering the high number of electrodes and the estimated extended distances between the probes and the measuring card, the method recommends using probes to convert the measured voltage to a current loop, making them suitable for measuring small voltage levels in the millivolt range.**

***Index Terms*—Electrical impedance tomography, active measuring probe, automatic measurements, phase shift measurement, correlation method, voltage and current measurements**

I. Introduction

Currently, many non-invasive methods exist for measuring an object with an unknown internal structure. The most common include X-ray, computed tomography, magnetic resonance imaging, ultrasound, positron emission tomography or single-photon emission tomography. These methods share common characteristics: they are expensive, often non-mobile, and require trained operators. For the general public, electro-impedance tomography remains less known. It measures the voltage between electrodes (usually 16) [1], [2]. Consecutively, the computer calculates conductivity matrices between the electrodes and displays them as a two-dimensional image. However, the method cannot be applied to everything because one electrode is always powered by current. The current levels can vary but usually are up to 10 mA. Hence, this method can be applicable to inanimate objects with small voltage or current sensitivity. The method requires no radioactive or magnetic material but only a current source and a measuring card. The real and imaginary impedance components must be measured to process the data. The impedance can be expressed by Equation (1), where **Z** is the complex impedance, R is the electrical resistance, and X the reactance:

$$\mathbf{Z} = R + jX. \tag{1}$$

II. Measuring real impedance component

The real impedance component is very easy to measure. All that is necessary is the value of the current flowing through the load and the voltage measured directly thereon. The resistance can be calculated using Ohm's law; equation (2):

$$R = \frac{U}{I}. \tag{2}$$

Whether it is used to measure amplitude, peak value or RMS value, the results (impedances) are the same; however, both signals must be measured in the same way [6]. A noise signal will affect the measurement. It is very likely at small amplitudes, i. e. that peak and amplitude measurements can be slightly misleading. Therefore, the RMS value has to be measured, or a filter has to be created to remove the noise. A prefab chip or a microprocessor can be used in the measuring card to measure the RMS value. Using an analogue filter is inefficient because one filter cannot work correctly at lower frequencies; thus, a bank of filters is required. Signal sampling and measuring the RMS voltage value in the microprocessor is the most suitable option here. The sampled signal is also required to calculate the imaginary impedance component. If there is relevant noise, the signal could still be filtered by a finite response filter with a variable cut-off frequency directly in the microprocessor.

III. Measuring the imaginary impedance component

The imaginary impedance is frequency-dependent. The correlation indicates that it is necessary to measure the phase shift between the current and voltage generated at the load. The most straightforward method for measuring phase shift is with zero-cross detector [6]. This requires a precision timer. For example, if we want to measure the phase shift of a 10 kHz signal with a resolution of 1 degree, the minimum frequency of the timer should be 7.2 MHz, see equation (3):

$$f_{\mathrm{TIMER}} = f_{\mathrm{signal}} \cdot 360 \cdot n = 10000 \cdot 360 \cdot 2 = 7.2\mathrm{MHz}, \tag{3}$$

where n is the number of samples for one degree (prevents output bouncing). The user must define the measured signal frequency for all methods, or the frequency must be calculated.

A phase shift measurement method with correlation was tested. The following section describes the enumeration of frequency using the disadvantages of optocouplers. The chapter VI discusses accuracy of this method.

One exploits that the signal is supposed to be electrically isolated from the computer to measure the frequency. Standard optocouplers (without any offset) operate only at positive voltages. If the input signal has negative polarity, the output is rectified. The concept utilises the rectified signal.

This feature is exploited to increase the dynamic range of the analogue-to-digital converter. As part of the signal processing phase, the microprocessor calculates the number of samples with zero value - the sampled signal is shown in Figure 1. The signal frequency is calculated using Equation (4), where n is the number of samples with zero value, and f_s is the sampling frequency:

$$f_{\text{signal}} = \frac{f_{\text{s}}}{2 \cdot n}. \tag{4}$$

The frequency is divided by two since the number of samples is number of samples in zero part (originally, the negative segment of the sinusoidal wave).

To get zero values unaffected by noise, the signal is first summed up with a slight offset around zero. After passing through the optocoupler, the offset is removed. This step ensures that there will be a small negative voltage at the AD converter input, represented as zeros.

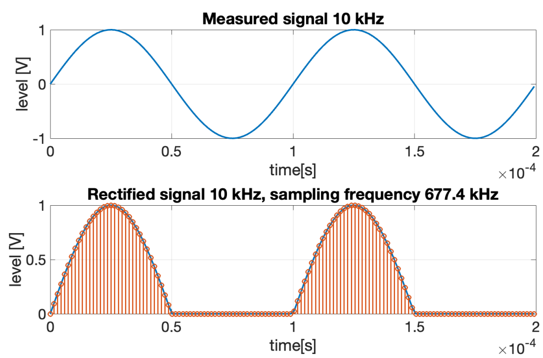


Fig. 1. Sampled rectified signal - frequency calculation

A. Phase shift - correlation method

If we imagine a signal with a millivolt level, additional noise shall be added thereto. The noise can stem from components such as resistors or external sources (transformers or power cables near the measurement card). The noise could cause an incorrect operation of the zero cross detector. Correlation-based methods have been developed. The device measures two signals: a time-dependent current waveform and the voltage between two probes. The signals are similar, only attenuated or shifted in time in some way. Therefore, the cross-correlation method was applied. Both signals are inserted into a cross-correlator. The algorithm was simulated in Matlab. Figure 2 illustrates the two simulated signals with Gaussian noise with a standard deviation of 1.2. The cross-correlation output is depicted in Figure 3.

The cross-correlation function's maximum value must be determined to evaluate the two signals' phase shifts. Subsequently, the phase shift is evaluated using Equation (5):

$$\varphi = i_{(f*g)\max} \cdot 360 \cdot \frac{f_{\text{signal}}}{f_{\text{sampling}}}, \tag{5}$$

where $i_{(f*g)\max}$ is the position of maximal value in the cross-correlation vector, f_{signal} is the measured frequency of one signal, and f_{sampling} is the sampling frequency of the analogue-to-digital converter. [5]

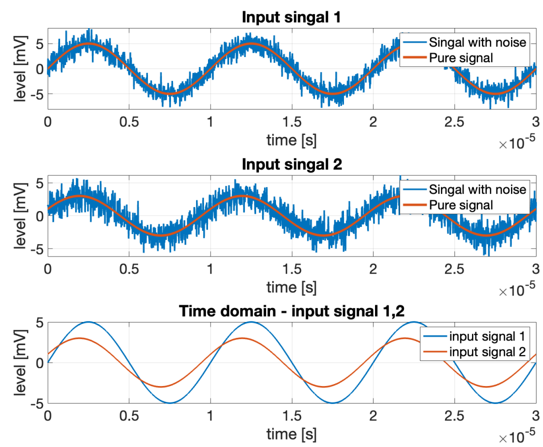


Fig. 2. Two time-shifted signals with an additional Gaussian noise ($\sigma = 1.2$) [3]

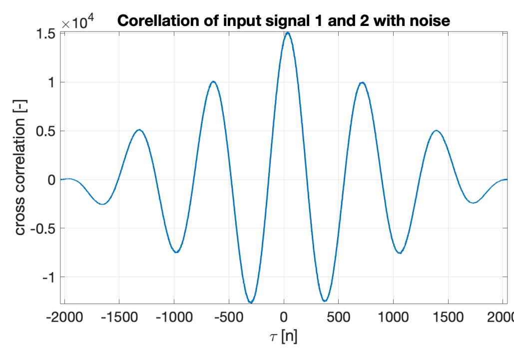


Fig. 3. Cross-correlation of the signals shown in Figure 2 [3]

The accuracy of this method was evaluated using Matlab. A signal with Gaussian noise with a standard deviation from the interval $\langle 0; 1.2 \rangle$ was generated in 10,000 experiments. The error of the calculated and defined phase shift was calculated for each trial. The average error for this interval was 5.4%. If the standard deviation interval is reduced to $\langle 0; 0.2 \rangle$, the average error changes to 2.06%, comparable with the error of the signal processed by the same method without additional noise. [3]

If a minimum of ten samples per period is considered, the same speed of timer (7.2 MHz) used for the zero-cross detector method can be evaluated for measuring the signal with a 720 kHz frequency. Applying the correlation method, the maximum frequency is 72 times higher than in the zero-cross method for the same speed of timer/sampler.

IV. MEASURING CARD DESIGN

The main advantage of this concept is its scalability. The entire measuring card consists of two types of printed circuit boards. The first is the microprocessor board (at the bottom of the device shown in Fig. 4). The board is equipped with two amplifiers with programmable gain and optocouplers for galvanic isolation between the computer and the measuring component. The board also provides power to all other boards and ensures communication with the computer. Figure 5 depicts a block diagram of such a board for voltage/current (differential) measurement.

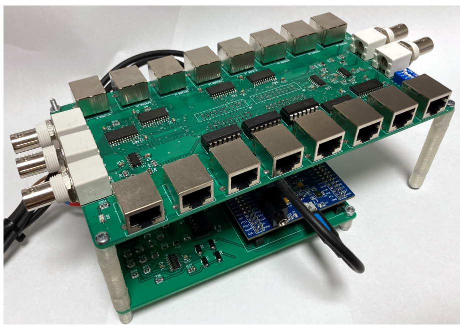


Fig. 4. Measuring card for electronic impedance tomography

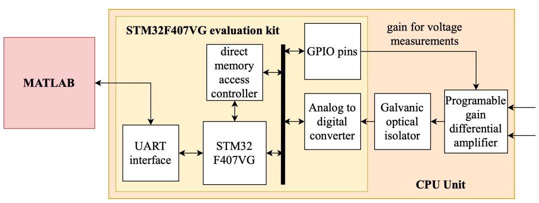


Fig. 5. Block diagram of the microprocessor board, simplified

Communication then takes place via a customised Matlab library (connected directly to EIDORS library) or via a simple terminal. The firmware defines commands for:

- Measurements between two electrodes,
- Measurements of the whole level,
- Multi-level measurements,
- Calibration function.

The second type of board is the measuring card. The assembly of this board depends on whether it is the first board (master) or an extension board (slave). In master mode, the board evaluates the signal processing performed on the active probes (discussed below in the text). The board also ensures switching the control signal to the probes (controls current delivery) and switching between measuring probes. In slave mode, the board is only responsible for the switching while the master evaluates the signal. The system's extensibility is ensured by addressing the board. This type of board can measure 16 probes. However, the board address can be set from 0 to 15. Thus, up to 256 electrodes could be connected by addressing. This can be used to measure more than one phantom or more planes.

Figure 6 shows a block diagram for collecting voltage from a single probe. The multiplexors are used as follows: two for controlling the signal, two for electrode addressing, and two for the attenuation signalisation. Control logic decides whether the attenuation should be switched.

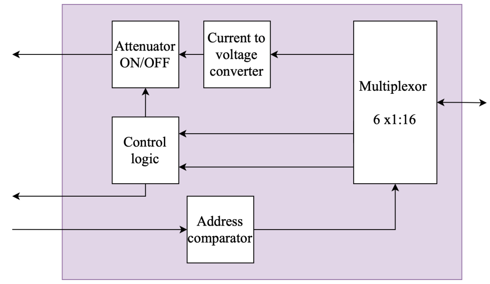


Fig. 6. Multiplexor block diagram - simplified

Multiplexors are of great importance because the current loop must be loaded with appropriate resistance. Therefore, the load resistance must be recalculated to the corresponding load after selecting the multiplexor. The remaining elements of the measurement chain are not strictly defined, and the selection can be variable. The same applies to programmable amplifiers, microprocessor or optocouplers.

V. ACTIVE PROBE DESIGN

The active probe is designed to eliminate all unwanted interferences and ensure high impedance at the measuring node. High impedance ensures measurement without affecting the flow of current through the object [6]. Its primary purpose is to convert the signal into a current loop. Doing so ensures that noise does not affect the signal, using induction from another signal or power supply. High impedance is provided by PI pads at inputs. Furthermore, using an attenuator, the probes protect against voltages higher than the operating voltage. The attenuator connects to the signal trace automatically if the input voltage is higher than the voltage set by the divider. Signal attenuation is transmitted to the measuring card via a single wire. The rationale being that to measure the voltage at two points, the signal must be attenuated at both electrodes or compensated in the microprocessor. If only one electrode signal is attenuated, errors in the evaluation may occur. Therefore, if only one signal is attenuated, the measuring card will attenuate the other. The attenuation is primarily used in power electrode measurements. The active probe is depicted in Figure 7; it is connected to the measuring card via a flat cable (without twisted pairs), consisting of RJ45 connectors. Figure 8 displays a simplified block diagram of an active probe.

The probes should also be able to switch power on and off or ground themselves to automate the measurement [2]. Two relays manage this function. The control signal from the measuring unit drives relays. Switching using unipolar transistors was tested in the simulations; however, the transient parasitic capacitance affects the power signal clarity. If the control signal is negative, the ground shall be connected to the node at the measurement point. If the signal is positive, the probe supplies current to the phantom. If the control signal oscillates around zero, both relays remain disconnected.

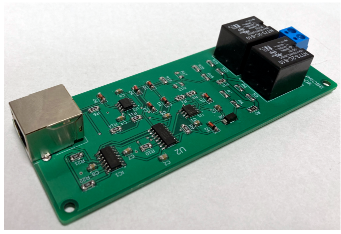


Fig. 7. Active probe

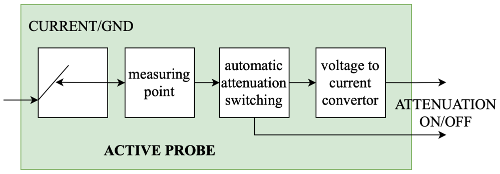


Fig. 8. Active probe block diagram, simplified

TABLE I
FREQUENCY CALCULATION ACCURACY

sample rate [kHz]		**677.4**	**1,354.8**
start frequency [kHz]	**stop frequency [kHz]**	**relative error [%]**	**relative error [%]**
0	12.2	5	2.5
12.2	26.8	10	5
26.8	44.2	15	10
44.2	60.4	20	10
60.4	78.1	25	15
78.1	84.6	30	15

VI. ESTIMATED PARAMETERS OF THE MEASURING CARD

The measurement error depends primarily on the electronic components used. The bandwidth is limited by the operational amplifiers used because the analogue signal processing does not include any other frequency-dependent constituents. The analogue-to-digital converter also limits the frequency. For the STM32F407VG (used in the prototype), the sampling frequency for the two synchronous captured channels is 677.4 kHz [4]. Considering the Nyquist theorem, the maximum signal speed is approximately 338 kHz.

For frequency enumeration, having at least ten samples for the period is preferable for better results. It limits the bandwidth to 67.74 kHz. This parameter can be improved by speeding up the sampling frequency. Speeding up the sampling frequency also improves the relative error of frequency enumeration. This is caused by the increasing number of samples needed for enumeration in the time domain. The values are calculated for 168 MHz core frequency. Table I shows the maximum estimated frequency measurement errors for two sample rates (systematic error).

The algorithm was changed to read the frequency from the user interface (known parameter); this error was eliminated. If the user provides an exact value for frequency in the command, enumeration will not be affected by this error. In contrast, if the user does not provide a frequency, the result will include a significant error. The bandwidth was enhanced to a minimum of 135.48 kHz (five samples to period). Changing the frequency enumeration to the zero cross detector method will block the processor for a while and requires more integrated circuits and inputs from the microprocessor. The phase shift calculation error is mentioned above.

The calibration function can adjust the measuring voltage and current level. The calibration function should measure voltage and current on a known resistive load. The algorithm is designed to measure the voltage on a resistor connected to two probes supplied with a known current. Consecutively, the microprocessor corrects all measured values with an absolute error calculated from the difference between uncalibrated and exact measurements, using a voltmeter and amperemeter. Such a procedure can only be performed between two probes, and the measurement offset should be the same for all probes (as long as the same components are used).

The calibration also corrects the phase measurement. By default, the phase cannot be zero for signals on resistive load. This is caused by the length of the measurement chain (for the current of approximately two amplifiers and voltage of at least four amplifiers). In calibration mode, the firmware quantifies the phase shift of the two signals on resistive load and then adds this constant to all the results.

VII. PERFORMED MEASUREMENTS

The concept is currently being debuged. However, the initial tests of the prototype yield positive results. The voltage transfer from the probe to the measuring card has been successful. It was tested on a harmonic signal, which appears free of significant distortion and has no additional superimposed noise.

Measurements using a microcontroller, under laboratory conditions, provide valid data. All algorithms applied are tested directly on a microcontroller and in Matlab.

VIII. ACKNOWLEDGEMENT

The preparation of this paper was assisted by the project No. FEKT-S-23-8425 of the BUT internal grant office.

REFERENCES

[1] DĚDKOVÁ, Jarmila. Metody rekonstrukce obrazu založené na elektrické impedanční tomografii: Image reconstruction methods based on electrical impedance tomography. Brno: VUTIUM, 2007. ISBN 978-80-214-3384-7.

[2] DUŠEK, Jan. Optimizing Image Reconstruction in Electrical Impedance Tomography. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Theoretical and Experimental Electrical Engineering, 2021, 123 p. Doctoral thesis. Advised by Assoc. Prof. Jan Mikulka

[3] ZIMNIOK, David. Analogově-digitální modul pro elektrickou impedanční tomografii. Brno: Brno University of Technology, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky, 2023, 75 p. Semestral Project. Advised by doc. Ing. Jan Mikulka, Ph.D.

[4] STMICROELECTRONICS, STM32F405xx STM32F407xx: Datasheet - production data [online]. In: . Rev9. 10/2020, p. 1-203 [Accessed 2023-03-27]. Available at: https://www.st.com/resource/en/datasheet/stm32f407vg.pdf

[5] Correlation for Time Delay Analysis. Panoradio SDR [online]. panoradio-sdr.de, 2019, 19.4.2019 [Accessed 2023-03-27]. Available at: https://panoradio-sdr.de/correlation-for-time-delay-analysis/

[6] CALLEGARO, Luca. Electrical Impedance: Principles, Measurement, and Applications. New York: CRC Press, 2012. ISBN 978-1-4398-4911-8.

System for measurement and control of pressure in vacuum apparatus

Sedlar J.
Faculty of Electrical Engineering and Communication,
DCI, Brno University of Technology
Technická 12, 616 00 Brno, Czechia
xsedla1x@vut.cz

Jurik K.
Faculty of Electrical Engineering and Communication,
DTEEE, Brno University of Technology
Technická 12, 616 00 Brno, Czechia
187537@vut.cz

***Abstract*—This paper summarizes the development and implementation of a device designed for controlling and operating a vacuum gas chamber. Multiple vacuum gauges, a gas pump and a valve can be connected to the device. Communication between the user and the system occurs via a large touch screen with an organized interface. The article describes the current state of development and outlines future improvement possibilities and potential applications.**

***Index Terms*—Vacuum control, pressure measurement, vacuum instrumentation**

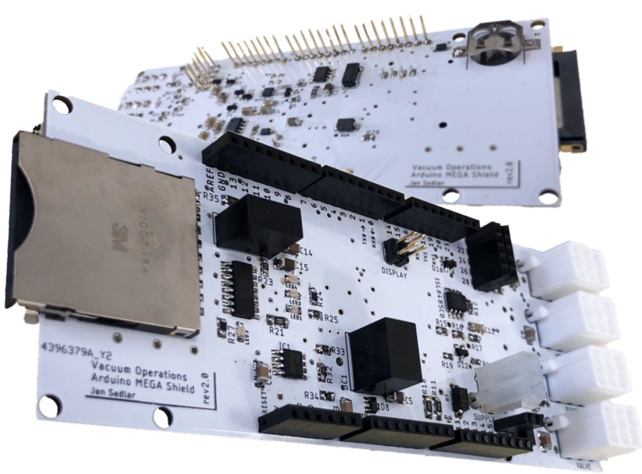


Fig. 1. Vacuum system controller custom PCB extension unit

I. Introduction

This article is aimed at describing the function of a system for controlling and operating a vacuum chamber. The main goal of such system is to monitor gas pressure in the chamber and to set the pressure according to user's requests.

The initial motivation for design of such a device was to enable laboratory research in low-pressure environments. Custom design of the pressure control unit potentially enables more possibilities and could provide the researches with greater flexibility by incorporating their requirements directly to the system, as opposed to relying on a commercial solution. Furthermore, the system's ability to offer functional control for different vacuum systems at a reduced cost could aid in offsetting the cost of research.

The heart of this system, which is the main focus of this article, is a printed circuit board (PCB) extension to an Arduino board that manages the communication with devices used for vacuum measurement, control and human-machine interface (HMI), further referred to as VSC (vacuum system controller).

The preparation of this paper was supported by the project No. FCH/FEKT-J-23-8282 of the BUT internal grant office.

II. Applications

The device described in this article enables easy integration of many broadly used vacuum operations equipment. The full potential of the system can be utilized for simultaneous gas pressure measurements and regulation but it is also possible to use just one of these functions if the application requires it. Although similar devices have been available on the market for many years, they might be difficult to obtain for small research teams or individuals due to the resources required for their acquisition. The purpose of the device described in this article is to provide a functional alternative with comparable quality to a professional-level system but with much greater accessibility.

III. System Description

The system is designed to be a simple, ready-to-use computer, with internal structure organized as shown in Fig. 2. This computer serves not only for communication and control of vacuum devices but also as their power source. All peripherals are connected via standardized ports typically used in similar applications. Such peripherals are up to two vacuum gauges used for gas pressure measurement, one vacuum pump removing gas particles from the chamber and one vacuum valve serving as a regulating device for controlled gas intakes.

The system's flexibility is increased by the fact that it can be used with various vacuum operation devices designed by *Pfeiffer Vacuum*. Users' requests and demands are inputted through a large Nextion touch screen panel embedded into the device body. The same screen is used to display various information regarding the current gas pressure and the general state of the vacuum system. To log this data, users can insert an SD card, which is subsequently used to store all important information for later evaluation of results.

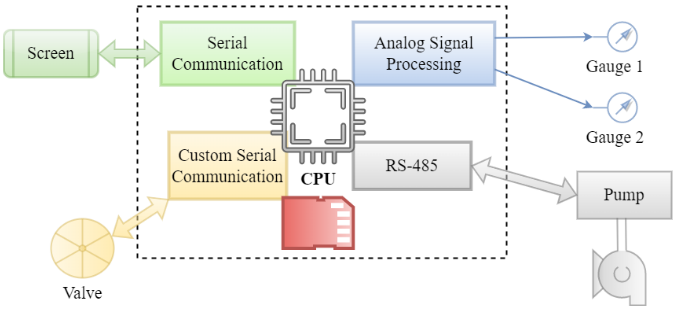


Fig. 2. System design diagram

As previously suggested, a custom PCB was designed and engineered to incorporate all essential circuitry required to implement the aforementioned features. The power-intensive peripherals were situated on the port side of the PCB, whereas the signal paths and processing components were positioned on other sections of the board, as demonstrated in Figure 3. A fully operational and assembled PCB extension unit can be observed in Figure 4.

IV. Communications

Different vacuum operations devices use varying communication protocols to send or receive data from the control unit. This had to be taken into account during the design of the VSC. While the pump and its controller are commanded via the standardized RS-485 interface [1] [2] [3], a more specific solution had to be implemented for the vacuum valve. The digitally controlled valve uses a half-duplex serial bus with atypical voltage levels defined for individual logical values and slow transmission speed (300 baud rate) [4]. The implemented solution is derived from the RS-232 communication protocol with its voltage level adjusted for this specific application. Both devices use a different set of instructions for their control.

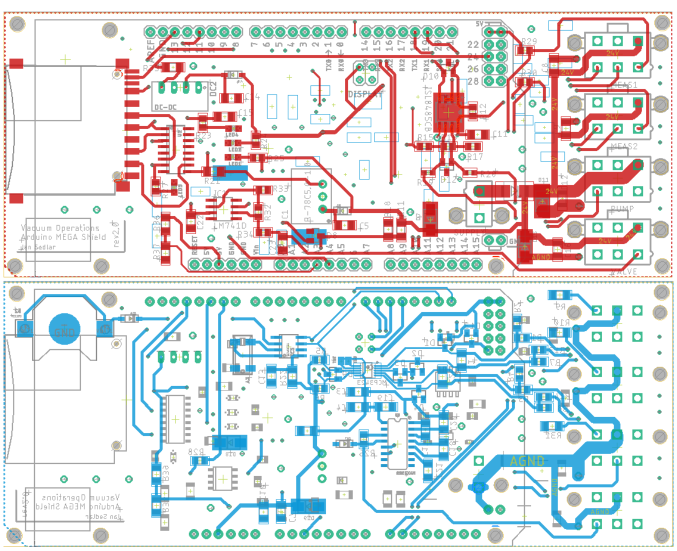


Fig. 3. 2-layer custom PCB extension design

The pump is commanded by *Pfeiffer Vacuum Protocol* messages [5], which are broadly used when more complex Pfeiffer vacuum products are concerned. The valve, on the other hand, implements a much simpler instruction list consisting of just a few commands enabling its basic control, settings and its current state determination [4]. Furthermore, to increase the flexibility of the VSC, a digital to analog conversion with a voltage step-up module was implemented for enabling the usage of a valve control via optional analog voltage line.

For the gas pressure acquisition, it was required to implement an analog to digital conversion, since many of the Pfeiffer Vacuum gauges represent the measured pressure as a voltage value [6] [7] [8]. This value is processed in an ADC (analog to digital converter) with high resolution of 18 bits providing no significant loss of precision during the conversion. ADC processes the data and sends the final output to the VSC via an I^2C interface. The same serial interface is used for acquiring or storing the data to the optionally inserted SD card.

For the internal communication with the Nextion HMI display a simple Arduino UART bus is used. The touch screen is also commanded by a standardized set of instructions and the device itself is programmed by a unique language.

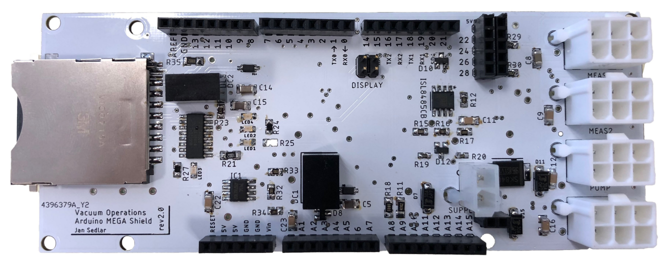


Fig. 4. Custom PCB extension unit top view

V. Power Supply

As previously mentioned, the VSC not only controls all connected peripherals but also acts as their main power source. A powerful source of up to 200 W had to be used, mainly because of the pump's large power consumption. The main power input circuitry of the VSC therefore accommodates multiple protections such as a basic protection against polarity reversal or against sudden peaks of input voltage. Additional pieces of protection hardware like Zener diodes, basic Schottky rectifiers or specialized serial communication lines protection diodes are located at every input leading into or output heading out of the VSC. The main power supply delivers a stable 24 V used to power all external peripherals, such as the pump and valve. Additionally, the internal circuitry operates at lower voltage levels, requiring two step-down converter circuits on the PCB, providing 5 V and 3.3 V, respectively.

VI. Central Processing Unit

In the current iteration of the project a widely used Arduino-based board is utilized as the main unit for processing. Specifically, an Arduino Mega with the ATMega 2560 chip is used which delivers enough computational power for servicing all incoming or outgoing requests [9]. The board was chosen for this application due to its multiple serial communication line modules and its easy expandability with additional circuitry, such as the custom-designed shield module for this particular application.

VII. Human-Machine Interface

The display's interface is designed with simplicity and ease-of-use in mind. The main screen (Fig. 5) displays all critical information related to the system's operation, while additional menus for data acquisition and storage, peripheral settings, and user presets can be accessed through multiple buttons located on the side of the main screen.

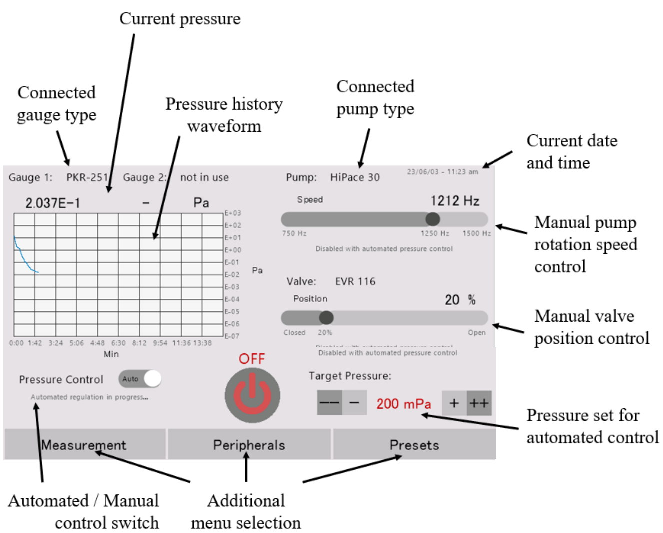


Fig. 5. Human-Machine Interface main screen

VIII. Software

The program running on the Arduino CPU must handle variety of tasks simultaneously, including the regulation of the gas pressure and communication with connected peripherals, accurate data acquisition or capturing and executing user's commands. While the current CPU is sufficient for these tasks, an upgrade to a more powerful or even multithreaded CPU shall be considered in the future when more demanding tasks are implemented.

With regards to the overall slow dynamic of the gas system, it is not necessary to have the parts of the algorithm servicing measurements and pump or valve control running in a loop at particularly high speeds. Data from the peripherals are acquired once every three seconds and the outgoing commands are sent at the same rate. On the other hand, user inputs must be handled in near real time, so the loop responsible for capturing these inputs repeats itself as frequently as possible to lower the response time of the touch screen to a bare minimum. The design of the Nextion displays represents a significant advantage – graphic tasks, such as rendering a new page or animations, are handled by the display's CPU and only important changes of the screen's state are sent over the serial line to be processed by the VSC.

At the time of writing this article a prototype of automated gas pressure regulation is being implemented, which will replace the currently used manual control of pump rotation speed or valve position. The purpose of such regulation is to achieve the gas pressure given by the user in the shortest time possible. In the upcoming iterations, the aim is to expand this feature by allowing the user to set the pressure progression in time. This would enable the creation of arbitrary gas pressure gradients, limited only by the vacuum hardware used in the application.

IX. Testing and Validation

The pressure values obtained as a voltage output from the VSC's 18-bit internal ADC were compared to those acquired with a laboratory-grade voltage measuring equipment of high precision. The error caused by the VSC is only given by the ADC's precision, as the gauge's accuracy cannot be influenced directly. The calculated error was found to be insignificant in comparison to the inherent error attributed to the gauges themselves, as anticipated.

The validation of the peripherals' correct functioning was done by evaluating the effect of the peripheral's state change (e.g. rotation speed) on the measured pressure values. Additionally, the confirmation of command acceptance by the peripherals was determined through the interpretation of acknowledgment response messages.

X. Future Improvements

Apart from already mentioned automated pressure control and its gradient that is currently being worked on, there are several other tasks that need to be addressed. Firstly, the VSC needs to be able to automatically identify the connected chamber system. This is important because the automated control requires precise knowledge of the gas chamber properties and must be readjusted after any changes to the system's topology. Secondly, a basic interface for remote access needs to be implemented to enhance the user experience and provide the ability to control ongoing experiments remotely. The vision here is to establish a basic internet server which is connected to the system and therefore enables the user to command or oversee currently ongoing system operation and not be directly present by the vacuum chamber itself. And finally, it is desirable to implement more connection possibilities for the usage of different vacuum modules which do not utilize the means of communication as described in IV. This will further broaden the device's flexibility and allow it to be used with a broader range of vacuum control systems without requiring a redesign of the VSC's internals.

XI. CONCLUSION

Although the current iteration may require several improvements to meet today's state-of-the-art devices, the first steps for a functional alternative to professional vacuum operations devices are already laid. The device is capable of communicating with a wide range of modules that are commonly used in vacuum chamber control, and can execute commands in accordance with the user's requests as input through the touch screen interface.

REFERENCES

[1] "Hipace 10 with tc 110 product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PMP03960.en.pdf. [online], [seen 10. 03. 2023].

[2] "Hipace 80 with tc 110 product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PMP03940.en.pdf. [online], [seen 10. 03. 2023].

[3] "Hipace 800 with tc 400 product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PMP04300.en.pdf. [online], [seen 10. 03. 2023].

[4] "Evr 116 motorized valve product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PFI39931A.en.pdf. [online], [seen 10. 03. 2023].

[5] "Pfeiffer vacuum-protokoll." Schnittstelle RS-232, RS-485.

[6] "Cold cathode vacuum gauge pkr251 product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PTR26000.en.pdf. [online], [seen 10. 03. 2023].

[7] "Cold cathode vacuum gauge ptr91 product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PTT03148310.en.pdf. [online], [seen 10. 03. 2023].

[8] "Pirani vacuum gauge tpr 270 product datasheet." https://static.pfeiffer-vacuum.com/productPdfs/PTR26770.en.pdf. [online], [seen 10. 03. 2023].

[9] ARDUINO, "Arduino.cc." https://arduino.cc. [online], [seen 10. 03. 2023].

Mobile application for electrophoretic gel image processing

Anna Mojžišová
Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xmojzi09@vutbr.cz

***Abstract*— This paper describes a mobile application that assists in the analysis of 1D gel electrophoresis images. It provides functions for image processing, band detection, lane segmentation, and molecular weight approximation. The designed methods were extensively tested using a dataset of diverse gel images. The application provides a convenient and portable tool for the analysis of electrophoresis images on-the-go.**

Keywords—1D gel electrophoresis, image analysis, quantitative analysis, mobile application

I. Introduction

One-dimensional gel electrophoresis is a widely used technique for separating proteins or nucleic acids in a sample based on their size and charge [1]. The sample is loaded onto a polyacrylamide gel, which acts as a molecular sieve [2, 3]. The gel is then placed in a buffer solution and an electric field is applied to the gel [4]. The molecules in the sample migrate through the gel at different rates based on their charge and size, resulting in banding patterns that can be visualized using staining techniques [5]. By comparing the banding pattern of the sample to that of known molecular weight markers loaded onto the gel, it is possible to perform a quantitative analysis of the molecular weights of separated molecules. The distance traveled by each band is proportional to the logarithm of the molecule's molecular weight, allowing the estimation of the molecular weight of the separated molecules [6]. This technique can be used for the identification and quantification of specific proteins or nucleic acids in a sample and can be used to monitor changes in protein expression or DNA mutations over time.

II. Methods

A. Grayscale conversion and image quality enhancement

First, the input image is converted to grayscale, and the mean pixel intensity is evaluated. This information is used to create a new grayscale image with dark bands and a lighter background. Numerous methods including histogram equalization [7, 8], gamma correction [7], piece-wise linear contrast adjustment [7, 9], and background thresholding [8][10] have been tested and evaluated. Out of these techniques, only gamma correction provided sufficient improvement in gel image quality.

Gamma correction is a technique used to adjust the brightness and contrast of an image by altering the gamma value of the image [11]. Gamma value was estimated by extracting pixel intensities from a representative region defined by a user containing all the sample lanes and minimum of additional background, computing the mean intensity of the pixels in the representative region, computing the logarithm of the pixel intensities and the logarithm of the mean intensity, fitting a line to the logarithmic data using linear regression, and extracting the slope of the fitted line, which is the estimated value of gamma [12].

B. Lane segmentation

The algorithm of lane detection was slightly modified from the method proposed by Škutková et al. and requires an input value of the number of lanes of the analyzed electrophoretic image [7]. The method involves two main steps. Firstly, the algorithm calculates the intensity mean value of each column in the upper third of the gel image and identifies peaks using findpeaks method to mark the position of the first pixel of each border. The algorithm automatically fills in the expected lane boundary if none is found. Secondly, the algorithm tracks the other pixels of each border line consecutively, comparing the intensity values of 3x3 pixels below the current pixel and selecting the column containing the pixel with the highest intensity. The tracking algorithm stops at the bottom of the gel. An example gel image with segmented lanes is shown in Fig.1.

C. Median filtering

In order to obtain a 1D signal representation of the intensity of pixels from 2D grayscale profiles representing individual sample lanes in horizontal orientation, a grayscale profile is firstly converted to a 1D signal by flattening it. Median filtering

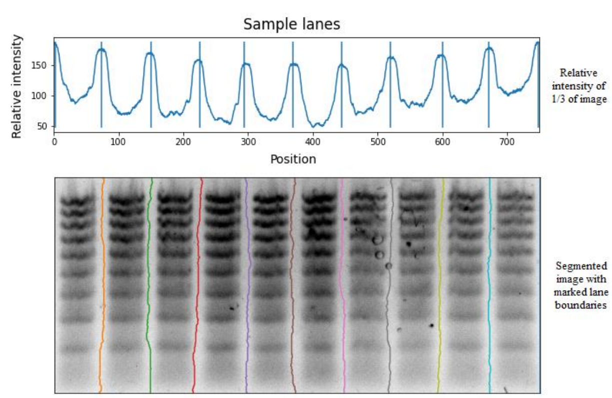


Fig. 1. Principle of gel image segmentation using tracking of lane boundaries

is then applied to the 1D signal using an appropriate window size, which is twice the width of the largest detected band [7].

D. Shading correction

In electrophoretic gels, the background can have a non-uniform shading due to factors such as uneven lighting or variations in the thickness of the gel [13]. By correcting the shading of the background, positions of bands can be more accurately measured and compared between different samples.

The steps involved in applying the local maxima method for envelope estimation include selecting an appropriate sliding window size, applying the window to the signal, computing the maximum value within the window, and repeating the process for each sample in the signal [14]. The result is then interpolated to obtain a curve covering the entire signal. The envelope is subsequently subtracted from the original signal, resulting in a corrected image with reduced background shading. The size of the sliding window is twice the size of the widest band in a signal. The window overlap is set as one quarter of the sliding window size and requires extension of both ends of the signal [7]. An example of input and output gel images is shown in Fig.2.

E. Band detection

The method involves analyzing a single-lane signal to detect bands. The first step is to invert the signal and then calculate the maximum amplitude. A prominence parameter is then calculated as a fraction of the maximum amplitude. This parameter is used to detect peaks in the signal using the find peaks function. The result is an array of peak indices representing the locations of the bands in the signal [15].

F. Molecular weights evaluation

This method involves creating a calibration curve to estimate the molecular weights of unknown protein or nucleic acid samples. The migration distance of each known molecular weight marker band and its corresponding molecular weight are obtained, and the calibration curve is created by interpolation of these data points. The molecular weights of the unknown bands in the sample lanes can be estimated by finding their corresponding distances in the 1D signal and then using the interpolation function to find their corresponding molecular weights [16].

III. Realization of Mobile Application

The structure of the application involves a client-side mobile application developed in Kotlin and a server-side application developed using Python and FastAPI. Simplified workflow is shown in Fig.3. The client-side application allows users to capture or upload 1D electrophoresis gel images and select the area of interest for further analysis. The user also provides additional information about the gel image, such as the number of sample lanes, the lane containing the molecular weight marker, and the type of molecular weight marker. This information is sent to the server-side application via a FastAPI endpoint. The server-side application retrieves the image data and performs image processing and analysis using Python-based code. This includes grayscale conversion, gamma correction, lane segmentation, median filtering, shading correction, and band detection.

To ensure that the pairing of marker bands and their corresponding molecular weights is accurate, a manual step has been introduced. This step is triggered if the number of detected bands does not correspond to the number of molecular weights values and requires user input to resolve any discrepancies. The client-side receives an image of horizontally oriented marker lane along with its 1D intensity profile showing detected band peaks and corresponding molecular weights. The user can manually reposition the darts to ensure that all molecular marker bands are correctly annotated. The adjusted marker lane is sent back to the server-side.

A calibration curve is created and used to estimate the molecular weights of all the bands in the gel. The final result is an image that shows the molecular weights of individual bands. The new image is sent back to the client-side application via the FastAPI endpoint and displayed to the user. The user can save the image to their device or send it by email.

To add a new molecular weight marker, the user can request the addition through the client-side application. The request is sent to the server-side application via a new endpoint, and the server-side application updates the SQL database by creating a new record in the MarkerTypes table for the new marker type. The SQL database has three tables: Images, Bands, and MarkerTypes. The Images table stores information about the electrophoresis gel images uploaded by the users. The Bands table stores information about the molecular weight bands detected in the images, including their position and the molecular weight of the band if known. The MarkerTypes table stores information about the available molecular weight marker types, including the name of the marker and values of molecular weights of its consecutive bands.

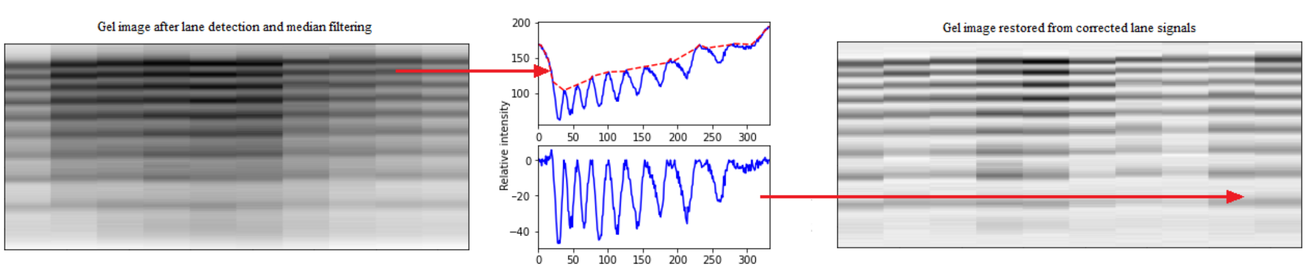


Fig. 2. The gel image after lane detection and median filtering (left), the signal of the lane with estimation of the signal envelope (middle), the reconstructed gel image with removed background shading (right).

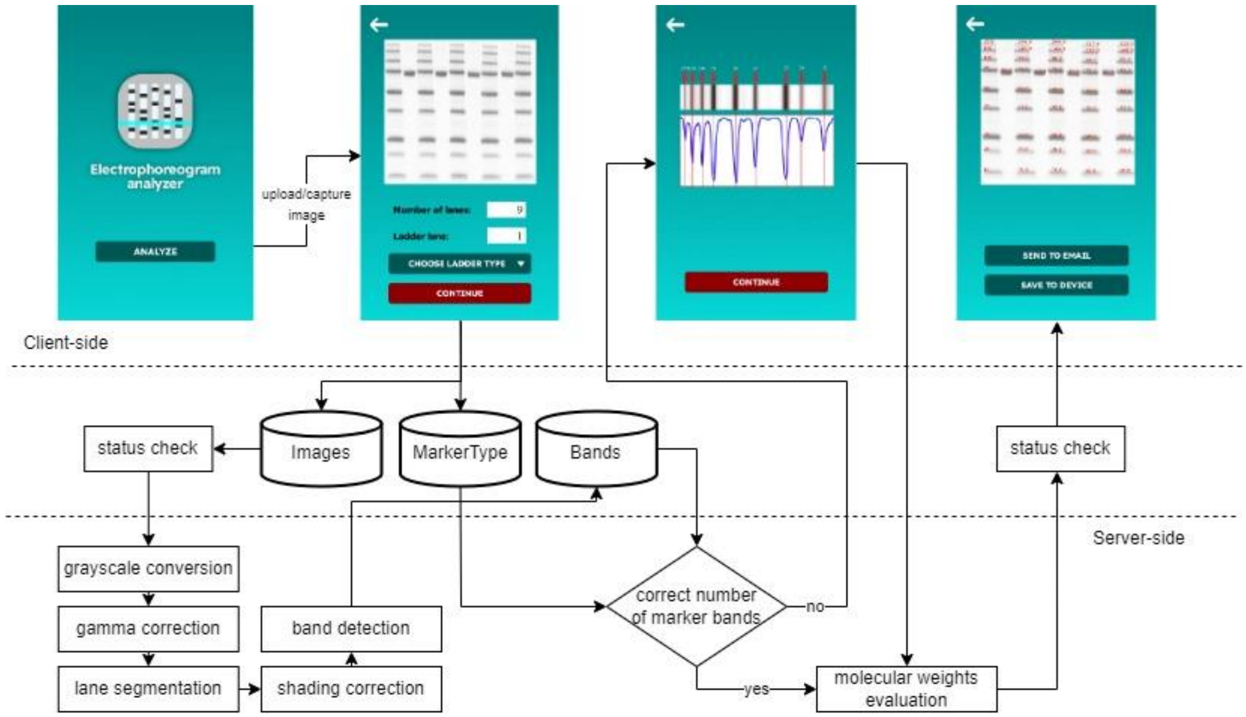


Fig. 3. Workflow of the mobile application.

The client-side and server-side components are connected through the use of five API endpoints, which are responsible for handling requests and responses between the mobile application and the server. Upload Endpoint receives a POST request containing a cropped gel image and additional information about the image. This endpoint stores the image and data in a SQL database on the server. Marker Endpoint allows users to add new molecular weight marker types to the application. It accepts a POST request with marker information and returns a response with the newly created marker ID. Analysis Endpoint processes the image and additional data stored in the SQL database by the upload endpoint. This endpoint uses Python code to perform the image analysis described in the previous chapter and generate the analyzed image. Download Endpoint returns the analyzed image to the mobile application in response to a GET request. The URL of this endpoint is used by the mobile application to retrieve the analyzed image from the server. Status Endpoint returns real-time updates on the progress of the image analysis. This endpoint can be used by the mobile application to monitor the progress of the analysis and estimate how much longer the analysis is expected to take. The actual analysis time may vary depending on the complexity of the image being. Typically, the analysis time for a single image can range from 5 to 30 seconds, with more complex images taking longer to process.

IV. RESULTS

A. Dataset

A dataset of electrophoretic gel images with known ground truth information obtained from laboratory measurements and external sources [17, 18, 19] was created. The ground truth information includes locations of all bands in each image, with a smaller subset of images including corresponding values of molecular weight. The images have been divided into several categories depending on their quality – standard, noise, non-uniform lane boundaries, blurred, and bubbles, to ensure that the program's algorithm can produce accurate results across different imaging conditions.

B. Lane segmentation

Accurate lane segmentation is crucial for the subsequent steps of band detection and molecular weight estimation. The lane segmentation method was applied to the dataset and the results were manually evaluated. The accuracy of the described method was calculated for different categories of image quality. The results are shown in TABLE I.

C. Band detection

Band detection functionality was evaluated by comparing the locations and sizes of the detected bands against the ground truth information for different image quality categories. Precision was calculated as the number of true positive detections divided by the total number of detected bands. Sensitivity was calculated as the number of true positive detections divided by the total number of ground truth bands. The results are shown in TABLE II.

D. Molecular weight approximation

To test the molecular weight approximation functionality of the program, a ground truth subset of images with known molecular weights was used. Approximated values of molecular weights were calculated and compared to the ground truth values. The accuracy of the method was evaluated using root mean squared error (RMSE), which measures the square root of the average squared difference between the predicted values and the true values and gives more weight to larger errors because of the squaring operation. The results are shown in TABLE III.

TABLE I. LANE SEGMENTATION RESULTS

Category	*No. of lanes tested*	*Accuracy [%]*
standard	112	100.0
noise	84	100.0
non-uniform	34	92.7
blurred	135	100.0
bubbles	47	100.0

TABLE II. BAND DETECTION RESULTS

Category	*No. of bands tested*	*Precision [%]*	*Sensitivity [%]*
standard	1038	99.8	99.6
noise	862	99.6	98.9
non-uniform	295	92.7	92.4
blurred	1264	95.0	94.8
bubbles	486	96.4	96.3

TABLE III. MOLECULAR WEIGHTS APPROXIMATION RESULTS

Category	*No. of bands tested*	*Relative RMSE [%]*
standard	148	3.4
blurred	184	5.4

V. DISCUSSION AND CONCLUSION

The lane segmentation method was found to be 100% accurate for most image categories, with the exception of images with non-uniform lane boundaries, which is primarily a problem of gel preparation and not the algorithm itself. The band detection algorithm performed well in high-quality images as well as images with noise. However, there is room for improvement in images with non-uniform lane boundaries and blurred lanes. Due to the limited availability of ground truth data, molecular weight approximation was not extensively tested.

In addition to the improvements to the existing methods, the application has the potential to be expanded for sample classification. Moreover, the application could include more extensive image quality improvement techniques to handle images with non-uniform boundaries and blurred lanes. Another potential feature is the adjustment of the mutual positions of the bands to allow more accurate detection and analysis of molecular weights. These potential advancements could greatly enhance the usefulness of the mobile application

In conclusion, the developed method for the analysis of molecular weights in gel electrophoresis images is accurate and efficient. It can analyze images on-the-go without requiring any specialized equipment or expertise, making it accessible to a wide range of researchers. While the method performs best on images without geometric distortions, it still provides highly accurate results for distorted images. Overall, this application has the potential to save researchers time and resources. Further development and refinement of the mobile application could lead to even greater accuracy and utility in the future.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisor, Ing. Helena Škutková, Ph.D., for her invaluable help, advice, and guidance throughout this project.

REFERENCES

[1] B. Alberts. Molecular biology of the cell, 5th ed. New York: Garland Science, 2008. ISBN 9780815341055.

[2] K. Wilson. Wilson and Walker's principles and techniques of biochemistry and molecular biology. Cambridge University Press, 2018. ISBN 9781316614761.

[3] J.L. Viovy. "Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms," in Reviews of Modern Physics, vol. 3, pp.813-872. 2000. ISSN 0034-6861.

[4] J. Sambrook, D.W. Russell. Molecular cloning: a laboratory manual, 3rd ed. New York: Cold spring harbor laboratory press, 2001. ISBN 9780879695767.

[5] S. Tseng, J.Y. Lin, J.P. Hsu. "Theoretical study of temperature influence on the electrophoresis of a pH-regulated polyelectrolyte," in Analytica Chimica Acta, issue 847, pp.80-89. 2014. ISSN 00032670.

[6] V.T.T. Lan, P.T.T. Loan, P.A.T. Duong, L.T. Thanh, N.T. Ha, T.B. Thuan. "Straightforward Procedure for Laboratory Production of DNA Ladder," in Journal of Nucleic Acids, pp.1-4. 2012. ISSN 2090-0201.

[7] H. Škutková, M. Vítek, S. Križková, R. Kizek, I. Provazník. "Preprocessing and classification of electrophoresis gel images using dynamic time warping," in International Journal of Electrochemical Science. 2013.

[8] N. Kaabouch, R.R. Schultz, B. Milavetz, L. Balakrishnan. "An analysis system for DNA gel electrophoresis images based on automatic thresholding and enhancement," in 2007 IEEE International Conference on Electro/Information Technology, pp. 26-31. IEEE, 2007. ISBN 9781424409402.

[9] Y. Labyed, N. Kaabouch, R.R. Schultz, B.B. Singh. "Automatic segmentation and band detection of protein images based on the standard deviation profile and its derivative," in 2007 IEEE International Conference on Electro/Information Technology, pp.577-582. IEEE, 2007. ISBN 9781424409402.

[10] L.G. Shapiro, G.C. Stockman. Computer vision. New Jersey: Prentice-Hall, 2001. ISBN 9780130307965.

[11] C.A. Poynton. Digital video and HDTV: Algorithms and interfaces. Morgan Kaufmann, pp.260. ISBN 1558607927.

[12] S. Rahman, M.M. Rahman, M. Abdullah-Al-Wadud. "An adaptive gamma correction for image enhancement," in J Image Video Proc, issue 35. 2016.

[13] R.B. Paranjape. "Fundamental Enhancement Techniques," in Biomedical Engineering, Handbook of Medical Imaging, Academic Press, pp.3-18. 2000. ISBN 9780120777907.

[14] A. Azqueta, K.B. Gutzkov, G. Brunborg, A.R. Collins. "Towards a more reliable comet assay: Optimising agarose concentration, unwinding time and electrophoresis conditions," in Mutation Research/Genetic Toxicology and Environmental Mutagenesis, vol.1, pp.41-45. 2011. ISSN 13835718.

[15] S.W. Smith. "Moving Average Filters," in Digital Signal Processing, pp.277-284. Newnes, 2003. ISBN 9780750674447.

[16] D.C. Tseng, Y.C. Lee. "Automatic band detection on pulsed-field electrophoresis image," in Pattern Anal Applic, pp.145-155. 2015.

[17] M.M. Muharram, M.S. Abdel-Kader. "Utilization of gel electrophoreses for the quantitative estimation of digestive enzyme papain," in Saudi Pharmaceutical Jpurnal, vol.25, pp.359-364. 2017. ISSN 13190164.

[18] T. Dvořáček. "Kompenzace geometrického zkreslení obrazu gelové elektroforézy." Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2015. 107s. Vedoucí diplomové práce Ing. Helena Škutková.

[19] RGP: 332 PCR-based genetic markers on rice chromosomes, National Institute of Agrobiological Sciences licensed under CC Attribution-Share Alike 2.1 Japan .

Age Estimation from Retinal Images: Different Image Preprocessing Approaches

1st Vojtěch Kadlec
Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication
Brno University of Technology, Czech Republic
Brno, Czech Republic
xkadle43@vutbr.cz

***Abstract*—Human age is considered an important biometric parameter that is often difficult to determine. Previous studies have shown that the non-specific general anatomical and physiological characteristics seen on fundus images are all likely signs of ageing. This paper focuses on age estimation from retinal images with different image preprocessing approaches together with proposed image detail enhancement method. Convolution neural network framework is based on the ResNet-34 architecture together with the Consistent Rank Logits algorithm estimating age as an ordinal variable. The best model achieved a mean absolute error of 3.47 years, outperforming existing models estimating age from retinal images.**

***Index Terms*—Deep learning, age estimation, retinal images, retinal image preprocessing**

I. INTRODUCTION

The advent of deep learning has enabled great advances in the field of medicine. It has been used to classify and segment medical images, discover new drugs or discover genetic abnormalities.

Recently, biological age has been investigated to reveal the predisposition to premature death. Several methods have already been developed to investigate this phenomenon. Basic biomarkers include DNA methylation status [1], in which a methyl group is linked to the pyrimidine cytosine and purine guanine. With the advancement of imaging methods, age estimation is also being made by this route from magnetic resonance images [2].

However, both principles have significant negatives in terms of financial or time costs that could be eliminated by retina in the future. It is the only organ of the human body that allows direct and non-invasive visualization of microvascular and neural tissue. Based on previous research, the similarity of anatomical and physiological features to vital organs has been established and the retina has come to be referred to as the window to the brain and heart [3] [4]. Thus, the retina may show signs of neurodegenerative and cardiovascular diseases, assuming a similarity with the molecular metabolic and morphological changes in the aforementioned organs and may play an important role in age prediction [5] [6].

This paper focuses on the different approaches of retinal image preprocessing for age estimation using a deep learning algorithm.

II. AGE CLASSIFICATION

Human age is considered to be an important biometric parameter, which is often difficult to determine at first sight. Complications arise due to complex ageing processes that are influenced by both internal factors, such as genes, and external factors, such as lifestyle or environment. In the field of machine learning, there are many methods based on different algorithms for determination and can be divided into 3 categories: classification [7], ranking methods [8] and regression [9].

III. CONSISTENT RANK LOGITS

One of the successfully applied method for age classification is Consistent Rank Logits (CORAL) with ordinal regression [10], which will be used also in this paper.

Ordinal regression is used in statistical analysis to predict ordinal variables whose values lie on an ordinal scale, where only the relative ordering between values can be considered significant.

For the purpose of the CORAL (Fig. 1), rank label y_i is extended into $K-1$ binary labels $y_i^{(1)}, ..., y_i^{(K-1)}$, where $y_i^{(k)\in\{0,1\}}$ indicates if y_i exceeds rank r_k (1). This indication is further denoted as a function $\mathbb{1}\{.\}$, which returns the value 1 if the inner condition is true, otherwise returns 0 [10].

$$y_i^{(k)} = \begin{cases} 1, & y_i > r_k \\ 0, & y_i \leq r_k \end{cases} \tag{1}$$

The predicted rank label r_q is obtained from the binary classifier $h(\mathbf{x}_i) = r_q$. Binary label vector q is then given by the sum of the k-th binary classifier $f_k(x_i \in \{0,1\}$ plus one (2).

$$q = 1 + \sum_{k=1}^{K} f_k(\mathbf{x}_i) \tag{2}$$

All predictions f_k reflect ordinal information together with rank-monotonicity, $f_1(\mathbf{x}_i) \geq f_2(\mathbf{x}_i) \geq ... \geq f_{K-1}(\mathbf{x}_i)$. For the age case, the rank-monotonic rule can be simply explained as follows: moving one ordinal value to a higher index must increase its actual non-ordinarily metric value, and the same must hold for moving in the opposite direction.

The weight parameters of the neural network, with the exception of the bias units of the output layer, are further

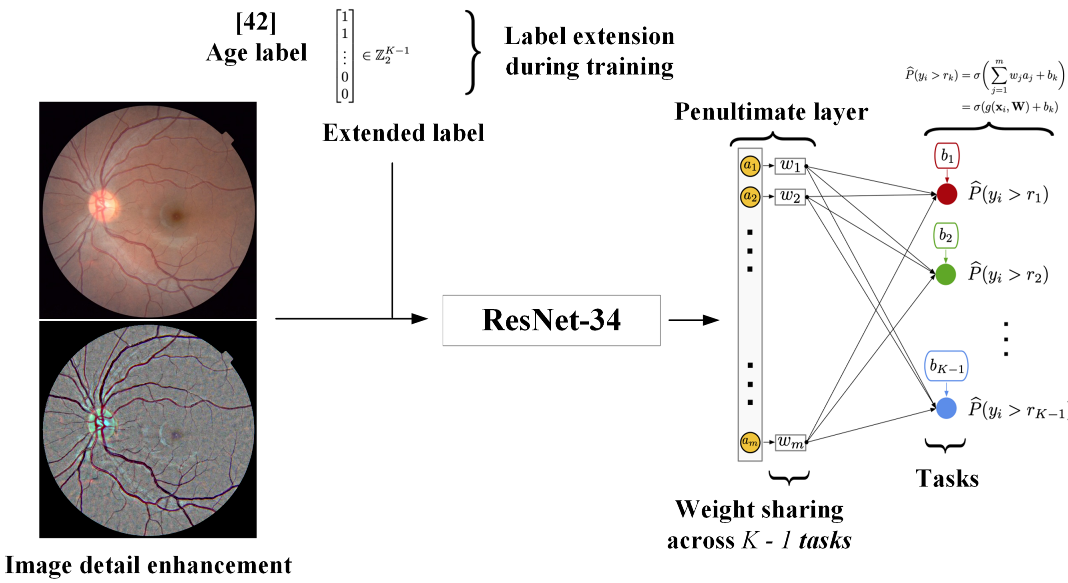


Fig. 1. The pipeline for age estimation. Binary labels are obtained from estimated probability via (5) and transformed to the age label using (2). [10]

denoted as $\mathbf{W}$. Penultimate layer with its output $g(\mathbf{x}_i, \mathbf{W})$, shares a single weight with all neurons of the output layer. $K-1$ independent bias units are then added to the output of the penultimate layer and an input $\{g(\mathbf{x}_i, \mathbf{W}) + b_k\}_{k=1}^{K-1}$ is created for the corresponding binary classifiers in the final layer. The predicted empirical probability for task k is defined in (3) with defined sigmoid function σ. [10]

$$\hat{P}(y_i^k = 1) = \sigma(g(\mathbf{x}_i, \mathbf{W}) + b_k). \tag{3}$$

The loss function of CORAL is the weighted cross-entropy of $K-1$ binary classifiers, denoted as

$$\begin{aligned} L(\mathbf{W}, \mathbf{b}) = -\sum_{i=1}^{N} \sum_{k=1}^{K} \lambda^k [\log\left(\sigma(g(\mathbf{x}_i, \mathbf{W}) + b_k)\right) y_i^{(k)} \\ + \log\left(1 - \sigma(g(\mathbf{x}_i, \mathbf{W}) + b_k)\right)(1 - y_i^{(k)})], \end{aligned} \tag{4}$$

where λ^k is the loss weight associated with the k-th classifier. Binary labels for rank prediction are obtained via

$$f_k(\mathbf{x}_i) = \mathbb{1}\{\hat{P}(y_i^{(k)} = 1) > 0.5\}. \tag{5}$$

IV. Image Processing

Preprocessing of retinal images plays an important role in their subsequent classification or segmentation. It is mainly used to remove unwanted artefacts or to highlight basic anatomical structures, especially arteries, veins and optic nerve head.

According to earlier research, the non-specific general anatomical and physiological characteristics visible on fundus images are all plausible indicators of ageing [4] [11]. Based on these theories, two methods were used to process the images.

A. Contrast Limited Adaptive Histogram Equalization

Contrast Limited Adaptive Histogram Equalization (CLAHE) is based on adaptive histogram equalization and eliminates its shortcomings in the form of noise amplification in homogeneous regions of the image [12]. This local operation is working with a predefined limit parameter. When a pixel intensity value exceeding the given parameter is found, its value is reduced. The part located above the limit is then spread among the values with lower intensity (Fig. 2).

B. Proposed Image Detail Enhancement

The method is based on the algorithm of a previously published paper [13]. The result of the extension removes the shortcomings of this particular algorithm, where image distortion occurs due to edge effects caused by the field of view of the retinal image.

In the initial stage of the approach, a binary retinal image mask $Y(x, y)$ is obtained by thresholding the input image $X(x, y)$ with the threshold T (6).

$$Y(x, y) = \begin{cases} 1, & X(x, y) \geq T \\ 0, & X(x, y) < T \end{cases} \tag{6}$$

By successive application of convolution and image interpolation, an extended mirrored image is created, i.e. an image without an unwanted background. Subsequently, by applying a convolution kernel filter in the form of a 2D Gaussian function (7) using the parameter σ determining the standard deviation of the Gaussian distribution, we obtain its smoothed form.

$$G(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{7}$$

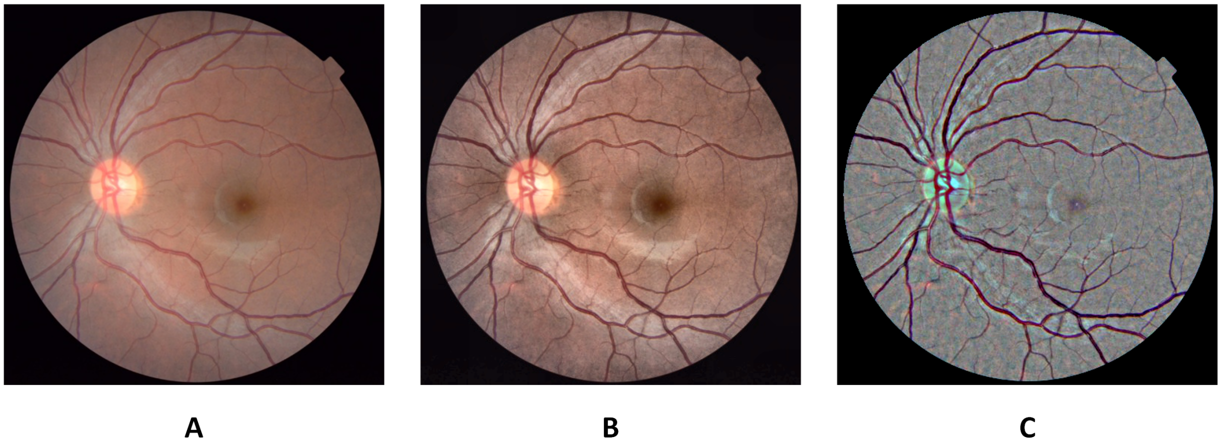


Fig. 2. (A) Original retinal image; (B) Retinal image after CLAHE application; (C) Retinal image after proposed detail enhancement method.

The resulting detail enhanced image (Fig. 2) is based on the application binary mask to a weighted version of the smoothed image together with the mirrored image.

V. Implementation

The available set of retinal images (N_{images}=7914) obtained using different types of retinal cameras was randomly split into three datasets - training (N_{images}=4656), validation (N_{images}=1956) and test (N_{images}=1302) with mean ages 51.07±11.25, 53.23±9.52 and 53.24±9.49 years respectively. All images belong to healthy subjects originating from Latin America.

One of the state-of-the-art CNN architecture - ResNet-34 was selected together with CORAL [14]. For the performance of the selected model, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were used, denoted as

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - h(\mathbf{x}_i)|, \tag{8}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - h(\mathbf{x}_i))^2}, \tag{9}$$

where y_i is the ground truth label of the i-th test example and $h(\mathbf{x}_i)$ is predicted label.

All images were cropped with retaining the main anatomical area of the fundus images, then resized to 256 × 256 after image detail enhancement. To prevent overfitting, various well-known augmentation methods were applied, random horizontal and vertical flip, random rotation of 0°-180°, random brightness with max delta 0.3 and random saturation with factor 0.1. After processing, the images were fed into the CORAL network, which returns a probability vector. Vector is then thresholded according to (5) and summed to the estimated age label.

The model was trained using an Adam optimizer. The learning rate was set to 0.0005. Model parameters were trained from scratch for 100 epochs. From those 100 epochs, the best model was selected via MAE performance on the validation set.

VI. Results

Deep learning model was developed using retinal images from 4656 subjects. The algorithm was trained on our dataset with three different principles, a network without including data preprocessing in terms of detail enhancement, a network with data preprocessing using the CLAHE method and a network using our proposed method for detail enhancement (DE). The results are shown in Tab. I. The best results achieved the network with the proposed image detail enhancement method with a MAE of 3.47 years and a RMSE of 4.65. The predicted age and chronological age have a fairly linear relationship between 40 and 60 years (Fig. 3). However, the prediction errors were higher for subjects under 45 and above 60 years. The predicted age of younger subjects showed to be overestimated and underestimated for older subject The worst performance was achieved by the algorithm without using any image detail enhancement. This fact confirms the hypothesis of the appearance of ageing markers in the anatomical structures of the retina.

TABLE I
Results of different methods of retinal image preprocessing

Method	MAE (yrs.)	RMSE
ResNet-34	4.21	5.89
ResNet-34 + CLAHE	3.78	5.12
ResNet-34 + DE	**3.47**	**4.65**

Our framework proved to work well in comparison to results available in other papers even though a much smaller dataset was available for network training. These results were also summarized in Tab. II, where our method outperformed other models.

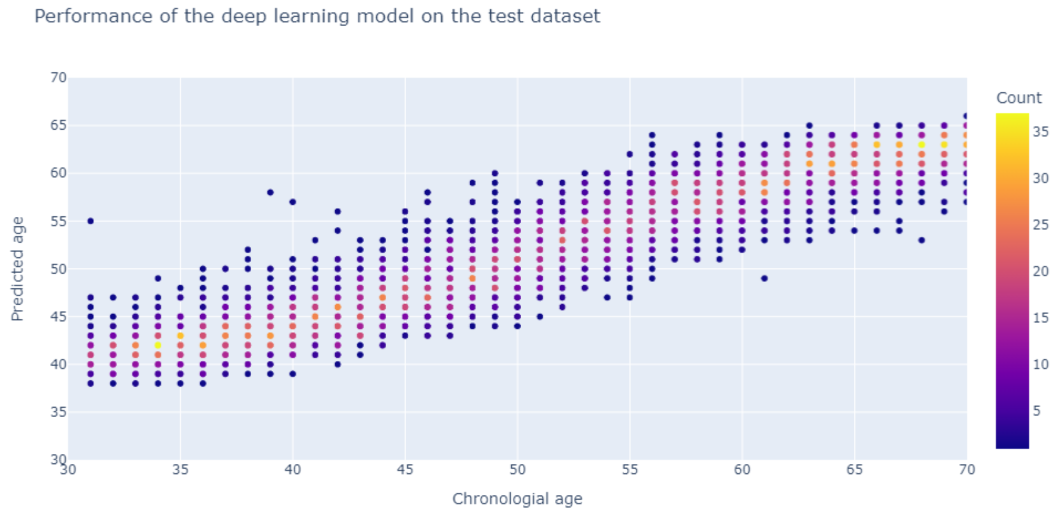


Fig. 3. Scatter plot showing performance of the deep learning model on the test dataset using a colour scale showing the most commonly predicted values for a given age.

TABLE II
MODELS AND RESULTS COMPARISON

Model	Source	Age range (yrs.)	MAE (yrs.)
ResNet-34 + DE	Ours	31-70	**3.47**
Xception	[15]	40-69	3.50
VGG-16	[13]	40-69	3.78
VGG-19	[16]	50-94	3.67

VII. Conclusion

Age estimation from retinal images has emerged as a promising technique for non-invasive assessment of human aging. The retina is a highly specialized and metabolically active tissue that undergoes structural and functional changes with age. This paper proposed the image detail enhancement method together with a CNN framework using ordinal regression for age estimation from retinal images. The developed deep learning algorithm can predict the human age with high accuracy. The overall MAE of the proposed method was lower than the recently reported models and the work thus points to the possibility of further research on retinal age. This may serve to reveal the predispositions to premature death.

Acknowledgment

This paper and the research behind it would not have been possible without the support of Aireen a.s., which provided the necessary retinal images. Additionally, I would like to express my heartfelt gratitude to my supervisor, Radim Kolář, for his support and guidance throughout this research project.

References

[1] Horvath, Steve, and Kenneth Raj. "DNA methylation-based biomarkers and the epigenetic clock theory of ageing." Nature Reviews Genetics 19.6 (2018): 371-384.

[2] Cole, James H., and Katja Franke. "Predicting age using neuroimaging: innovative brain ageing biomarkers." Trends in neurosciences 40.12 (2017): 681-690.

[3] Flammer, Josef, et al. "The eye and the heart." European heart journal 34.17 (2013): 1270-1278.

[4] London, Anat, Inbal Benhar, and Michal Schwartz. "The retina as a window to the brain—from eye research to CNS disorders." Nature Reviews Neurology 9.1 (2013): 44-53.

[5] Wong, Tien Yin, et al. "Retinal microvascular abnormalities and incident stroke: the Atherosclerosis Risk in Communities Study." The Lancet 358.9288 (2001): 1134-1140.

[6] La Morgia, Chiara, et al. "Retinal ganglion cells and circadian rhythms in Alzheimer's disease, Parkinson's disease, and beyond." Frontiers in neurology 8 (2017): 162.

[7] Pan, Hongyu, et al. "Mean-variance loss for deep age estimation from a face." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[8] Chen, Shixing, et al. "Using ranking-CNN for age estimation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[9] Fu, Yun, and Thomas S. Huang. "Human age estimation with regression on discriminative aging manifold." IEEE Transactions on Multimedia 10.4 (2008): 578-584.

[10] Cao, Wenzhi, Vahid Mirjalili, and Sebastian Raschka. "Rank consistent ordinal regression for neural networks with application to age estimation." Pattern Recognition Letters 140 (2020): 325-331.

[11] Cheung, Carol Yim-lui, et al. "Imaging retina to study dementia and stroke." Progress in retinal and eye research 57 (2017): 89-107.

[12] Setiawan, Agung W., et al. "Color retinal image enhancement using CLAHE." International conference on ICT for smart society. IEEE, 2013.

[13] Rim, Tyler Hyungtaek, et al. "Prediction of systemic biomarkers from retinal photographs: development and validation of deep-learning algorithms." The Lancet Digital Health 2.10 (2020): e526-e536.

[14] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[15] Zhu, Zhuoting, et al. "Retinal age gap as a predictive biomarker for mortality risk." British Journal of Ophthalmology (2022).

[16] Liu, Chi, et al. "Biological age estimated from retinal imaging: a novel biomarker of aging." Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22. Springer International Publishing, 2019.

ABB v Brně
Lepší svět začíná s vámi

Jsme součástí globálního technologického lídra působícího v oblastí 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

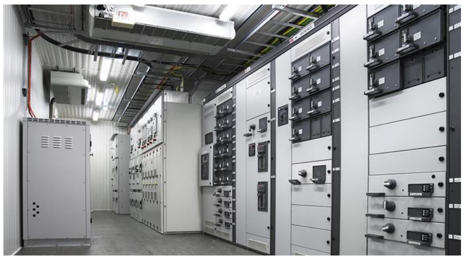
Virtuální prohlídka

Naše jednotky

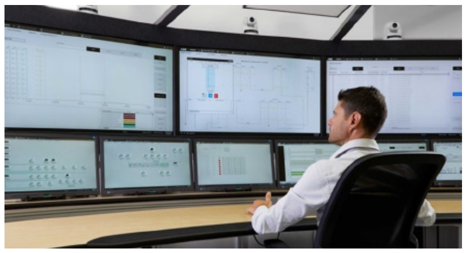
Rozváděče vysokého napětí
Výroba rozváděčů vysokého napětí se v našem závodě těší dlouhé tradici a pyšní se statusem největšího závodu svého druhu v Evropě. Co do počtu zaměstnanců, tak i plochy výroby, se jedná o největši jednotku v rámci brněnského závodu. Své produkty dodává do přibližně 100 zemí 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.

Implementation of a deep learning model for vertebral segmentation in CT data

Lenka Blažková
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
xblazk01@vutbr.cz

Michal Nohel
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
xnohel04@vutbr.cz

***Abstract*—This paper deals with the problem of vertebral segmentation in CT data with the use of deep learning approaches. Automatic segmentation of vertebrae is a very complex issue and would simplify the work of radiologists and doctors. The paper is focused on one of the models published and submitted to the Large Scale Vertebrae Segmentation Challenge (VerSe) in 2020 from C. Payer et al. – Improving Coarse to Fine Vertebrae Localisation and Segmentation with SpatialConfiguration-Net and U-Net and its implementation and modification. The model is evaluated on the corresponding public and hidden dataset. Its modification shows an improvement of the results in comparison with the published results, a mean Dice score improved from 0.9165 to 0.9302 on the public dataset and from 0.8971 to 0.9264 on the hidden dataset.**

***Index Terms*—deep learning, convolutional neural networks, vertebrae segmentation, segmentation, spine, vertebra, CT, computed tomography**

I. Introduction

The spine is the main part of the human musculoskeletal system, therefore, it is very important as a support of the human body and also as a protection of the spinal cord. Due to these important functions, spine analysis has a significant role in medicine.

Automatic segmentation of vertebrae would simplify the work of radiologists and doctors, e.g. with diagnosing various diseases, therapy planning or spine surgery planning. However, all structures of the spine are very variable in shape even without serious pathologies and image composition itself also quite varies, thus automatic segmentation is a complex challenge.

This paper deals with an implementation of a model for the automatic segmentation of vertebrae using 3D data obtained from computed tomography (CT) with the use of deep learning, more precisely convolution neural networks (CNNs). For vertebrae segmentation can be used different types of CNNs. One of the two most used CNNs is U-Net [1] and its modification, e.g. Attention U-Net [2] and ResUNet++ [3]. The second most used are ResNet [4] and Res2Net [5] and their modifications, e.g. PSPNet [6] and DenseASPP [7].

The authors of the chosen segmentation model Improving Coarse to Fine Vertebrae Localisation and Segmentation with SpatialConfiguration-Net and U-Net are C. Payer et al. [8] and for the segmentation, they use a modified U-Net. This model was selected because it was publicly available, published within the Large Scale Vertebrae Segmentation Challenge (VerSe) [9]–[11] which was organised in conjunction with the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), and also because it was one of the top-performing algorithms in this challenge.

II. Model description

The chosen model was submitted in VerSe'19, improved afterwards, submitted again within VerSe'20 and published. The model version from VerSe'20 is described below and its code can be accessed at: https://github.com/christianpayer/MedicalDataAugmentationTool-VerSe.

The algorithm is composed of three main steps – three CNNs. The first step after preprocessing is a localization of the spine in the CT volume. The second step is a localization of individual vertebrae and their identification. The third step is a binary segmentation itself which is followed by final postprocessing.

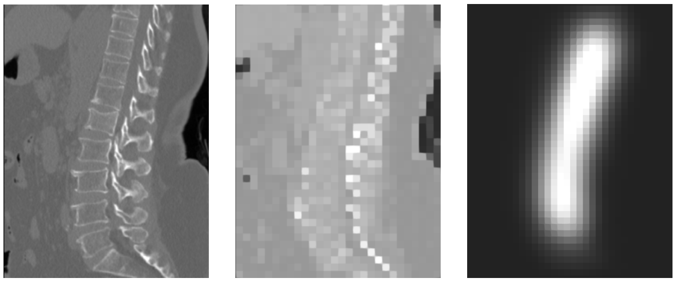
A. Preprocessing

The model uses composed preprocessing. Data preprocessing consists of three successive steps. The first step is a reorientation of the CT volume to RAI direction (Right-Anterior-Inferior). The next step is a Gaussian smoothing with $\sigma = 0.75$ to reduce noise and unfortunately also details. The last step is a clamping of the volume intensity values.

B. Spine localization

The first step in the volume analysis is finding the actual localization of the spine in the volume to reduce the region of interest. Firstly, the preprocessed volume is retrieved (Fig. 1 (a)) and resampled to uniform voxels of a size $8\times8\times8$ mm (Fig. 1 (b)) which means that a maximal size of the input can be $512\times512\times1024$ mm and after resampling $64\times64\times128$ voxels.

Then the volume enters the first CNN. The spine localization is performed by a modified U-Net with average-pooling and linear upsampling. Its architecture is based on five levels consisting of convolution layers with a kernel size $3\times3\times3$, 96 filters, zero padding and a Leaky ReLU activation function. The output of the network is a Gaussian heatmap and coordinates of the predicted spine localization (Fig. 1 (c)).



(a) (b) (c)

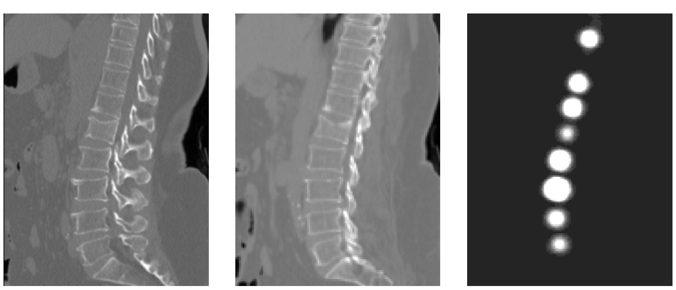
Fig. 1. An example of the spine localization steps: (a) input data, (b) resampled data and (c) final Gaussian heatmap of the spine localization

C. Vertebrae localization and identification

The next step in the algorithm is the localization and identification of individual vertebrae using the second CNN. Again the preprocessed volume (Fig. 2 (a)) is resampled, but this time to voxels of a size 2×2×2 mm (Fig. 2 (b)) and because of that the network input volume can have a size of 96×96×128 voxels at most and the volume before resampling can be 192×192×256 mm maximum.

If the input volume is greater than this size limit, it is cropped into subvolumes overlapping by 96 slices. These subvolumes are then processed independently and their results are merged at the end.

For vertebrae localization is used a SpatialConfiguration-Net (SC-Net) which consists of two blocks. The first block partially resembles a four-level U-Net where each level has two convolution layers with a kernel size 3×3×3, 96 filters and a Leaky ReLU activation function. The second block operates with a quarter of the volume resolution and has four convolution layers with kernels 7×7×7. The outputs of these two blocks are connected into Gaussian heatmaps of individual vertebrae. The individual heatmaps are then combined into one final Gaussian heatmap (Fig. 2 (c)) using thresholding and certain rules which are corresponding to the structure of the human spine, e.g. determination of the most upper vertebrae, the most lower vertebrae and the ones in between while all vertebrae must be spaced at least 12.5 mm and at most 50 mm apart.



(a) (b) (c)

Fig. 2. An example of the vertebrae localization steps: (a) input data, (b) resampled data and (c) final Gaussian heatmap of the vertebrae localization

D. Vertebrae segmentation

The final step is the actual binary segmentation of vertebrae. Each localized vertebra from the previous network is segmented individually. The preprocessed volume (Fig. 3 (a)) is cropped into subvolumes for each vertebra and resampled to the finest resolution of voxels of a size 1×1×1 mm (Fig. 3 (b)), thus a maximum input size into the CNN can be 128×128×96 voxels, 128×128×96 mm.

The last CNN is again a U-Net with a similar configuration as the one from the spine localization. However, the network outputs are segmentation masks of every vertebra separately (Fig. 3 (c)). These masks are then resampled into their correct position within the spine (Fig. 3 (d)) and finally combined into a single multi-label segmentation mask (Fig. 3 (e)).

E. Postprocessing

The final postprocessing of the resulting segmentation mask consists of transformation and reorientation of the created segmentation mask into the original orientation of the CT volume.

III. Data and metrics

Within VerSe'19 and VerSe'20 their datasets [9]–[11] were also created and published. The chosen model used the VerSe'20 dataset, consisting of three folds – training, validation (PUBLIC) and test (HIDDEN) data. The entire VerSe'20 dataset consists of 319 CT scans from 300 patients and 4141 labeled and segmented vertebrae. The CT scans were acquired using CT scanners from different manufacturers and with different scan settings. The scans include physiological cervical, thoracic and/or lumbar spine or with various pathologies. Besides the actual scans, the dataset contains two different types of annotations: voxel labels as binary segmentation masks and coordinates of vertebrae centroids, both created by automatic computer algorithms and several human experts.

As an evaluation metric for the segmentation masks is used a Dice score ($Dice$), more precisely its modification as a mean Dice score (1).

$$Dice(P,T) = \frac{1}{N}\sum_{i=1}^{N}\frac{2\,|P_i \cap T_i|}{|P_i| + |T_i|}, \tag{1}$$

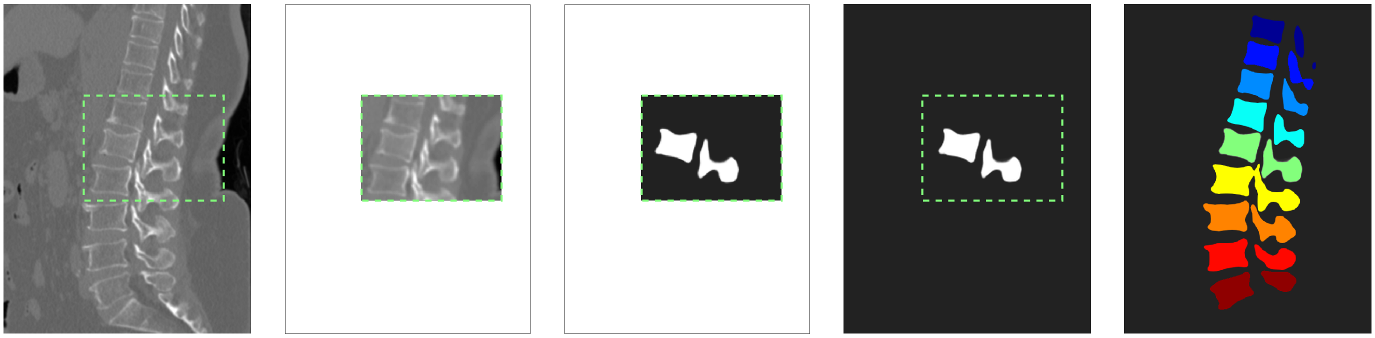
where P represents predicted voxel values, T correct voxel values (ground truth) and i is an index of N^{th} vertebrae. Final Dice score has no unit [–] and ranges between 0 and 1 (1 being the absolute match and 0 being no match).

IV. Results and discussion

The described model was implemented in two ways. Firstly, it was implemented with Docker image, since it was a part of the publicly available model. However, the outputs of this method are just the final segmentation masks without any debugging files or images. Therefore, the model was implemented in Python[1] and PyCharm[2] which provided more

[1]https://www.python.org/

[2]https://www.jetbrains.com/pycharm/



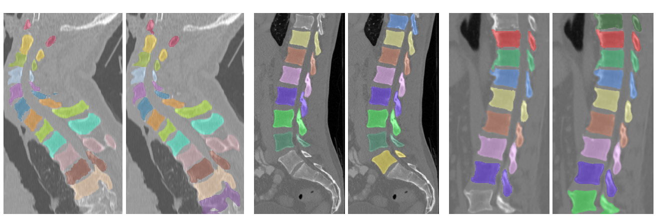
(a) (b) (c) (d) (e)

Fig. 3. An example of the vertebrae segmentation steps: (a) input data, (b) resampled and cropped data of one vertebra, (c) binary segmentation mask of one vertebra, (d) resampled binary segmentation mask of one vertebra into position within the spine and (e) final multi-label segmentation mask of all vertebrae

detailed understanding of the model, possibility to make changes in the algorithm and to save individual steps from each of the networks.

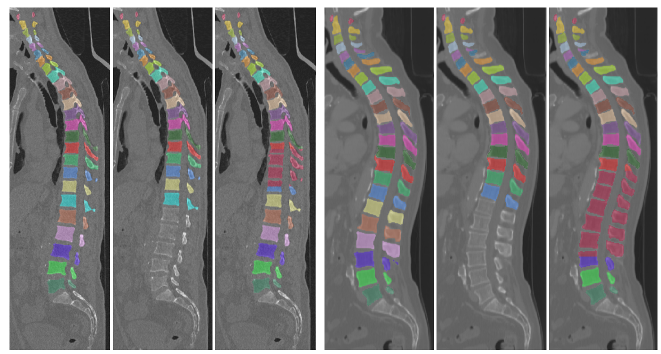
Paper [8] featured results of this algorithm with a mean Dice score of 0.9165 and a median of 0.9572 on the public data and a mean Dice score of 0.8971 and a median of 0.9565 on the hidden data. However, when the model was launched from GitHub and its outputs were evaluated, its mean Dice score was 0.8978 and its median was 0.9330 on the public data and 0.8918 and 0.9211 on the hidden data.

After further analysis of the outputs of the model, it was determined that some of the segmentation masks were created inaccurately in comparison with the ground truth masks. There were three types of segmentation defects: extra segmented vertebrae than in the ground truth (Fig. 4 (b)), unsegmented (missing) vertebrae (Fig. 5 (b)) or poorly segmented vertebrae (Fig. 6 (b)). The unsegmented vertebrae in masks were caused by incomplete vertebrae localization in the second CNN.



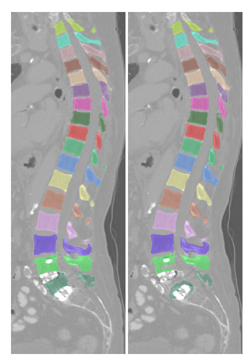
(a1) (b1) (a2) (b2) (a3) (b3)

Fig. 4. Examples of segmentation masks with extra segmented vertebrae together with their mean Dice score: (a) ground truth and (b) segmentation mask from the launched model of subject *verse352* (0.8804), *verse609* (0.8206) and *verse754* (0.8355)



(a1) (b1) (c1) (a2) (b2) (c2)

Fig. 5. Examples of segmentation masks with unsegmented (missing) vertebrae and their corrections together with their mean Dice score: (a) ground truth, (b) segmentation mask from the launched model and (c) segmentation mask from the modified model of subject *verse602* (0.7251; 0.9562) and *verse759* (0.6760; 0.9459)



(a) (b)

Fig. 6. An example of segmentation masks with poorly segmented vertebrae together with its mean Dice score: (a) ground truth and (b) segmentation mask from the launched model of subject *verse705* (0.8991)

Because of these inaccuracies, an extension of the model was performed. The model extension consists of cropping the CT volumes with an overlap of 96 slices, processing them apart and then connecting the final segmentation masks. This modification in the majority of cases provided better segmentation results in the problematic volumes, see Fig. 5 (c). After the modification, the results of the mean Dice score on the public data changed from 0.8978 to 0.9302, its standard deviation from 0.0884 to 0.0315 and its median from 0.9130 to 0.9396. Changes in the mean Dice score on the hidden data were from 0.8918 to 0.9264, its standard deviation from 0.0881 to 0.0307 and its median from 0.9211 to 0.9378.

In Fig. 7, you can see outliers (red +) detected using a Dice score and their improvement after the modification. Some of these outliers are shown in Fig. 5. All results (published, from the launched model and from the modified model) can be seen in Table I and in Table II.

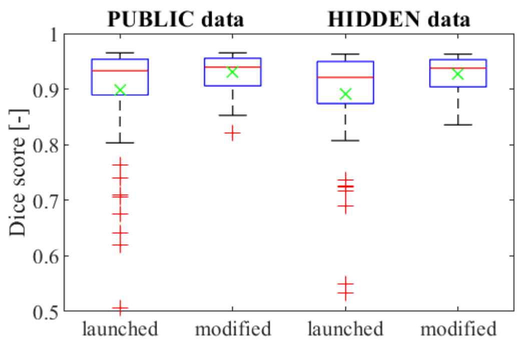


Fig. 7. Box and whisker plots of Dice score distribution from the launched and modified model on the public and hidden data: mean (green ×), median (red line), 25^{th}–75^{th} percentile (blue box), extreme values (black whiskers) and individual outliers (red +)

TABLE I
FINAL RESULTS ON PUBLIC DATA

Dice score	*published*	*launched*	*modified*
mean value	0.9165	0.8978	**0.9302**
std value	–	0.0884	**0.0315**
median value	0.9572	0.9330	0.9396
maximal value	–	0.9650	0.9650
minimal value	–	0.5072	0.8206

TABLE II
FINAL RESULTS ON HIDDEN DATA

Dice score	*published*	*launched*	*modified*
mean value	0.8971	0.8918	**0.9264**
std value	–	0.0881	**0.0307**
median value	0.9565	0.9211	0.9378
maximal value	–	0.9630	0.9630
minimal value	–	0.5327	0.8355

V. CONCLUSION

This paper aimed to implement and evaluate one of the deep learning vertebral segmentation models which would be a good help, e.g. with diagnosing certain spine diseases, therapy or surgery planning which would simplify radiologists' and doctors' work.

The chosen model was created by [8] and uses a modified U-Net for the vertebrae segmentation. In comparison with other algorithms that were submitted to VerSe'20, the chosen algorithm performed in the segmentation task as the second best. The model was tested on the two independent datasets (80 and 73 subjects) consisting of CT volumes and their annotations with various pathologies and different parts of the spine with a mean Dice score of 0.9165 and 0.8971 on the public and hidden data.

However, the published results in [9] did not correspond with the acquired ones from the model implementation. The published Dice score was lower. Therefore, the model was modified to perform on the inaccurately segmented subjects by processing the volumes in individually cropped subvolumes which gave even better results than the ones published. The mean Dice score was 0.9302 with a standard deviation of 0.0315 on the public part of the VerSe'20 dataset and 0.9264 on the hidden part with a standard deviation of 0.0307.

ACKNOWLEDGMENT

I would like to thank my supervisor Michal Nohel for guidance through this project, the authors of Verse'20 [9]–[11] for providing the VerSe dataset and publishing results of the challenge and also the authors of [8] for publishing their vertebral segmentation model.

REFERENCES

[1] O. Ronneberger, P. Fischer and, T. Brox, Eds., "U-Net: Convolutional Networks for Biomedical Image Segmentation", *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III*, no. 9351, pp. 234-241, 2015.

[2] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa2, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas".

[3] D. Jha, P. H. Smedsrud, M. A. Riegler, D. Johansen, T. D. Lange, P. Halvorsen, and H. D. Johansen, "ResUNet++: An Advanced Architecture for Medical Image Segmentation", *2019 IEEE International Symposium on Multimedia (ISM)*, pp. 225-2255, 2019.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.

[5] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, and P. Torr, "Res2Net: A New Multi-Scale Backbone Architecture", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 2, pp. 652-662, 2021.

[6] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid Scene Parsing Network", *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6230-6239, 2017.

[7] M. Yang, K. Yu, C. Zhang, Z. Li, and K. Yang, "DenseASPP for Semantic Segmentation in Street Scenes", *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3684-3692, 2018.

[8] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Coarse to Fine Vertebrae Localization and Segmentation with SpatialConfiguration-Net and U-Net", *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020)*, vol. 5, pp. 124-133, 2020.

[9] A. Sekuboyina, M. E. Husseini, A. Bayat, M. Löffler, H. Liebl, H. Li, G. Tetteh, J. Kukačka, C. Payer, D. Štern, M. Urschler, M. Chen, D. Cheng, N. Lessmann, Y. Hu, T. Wang, D. Yang, D. Xu, F. Ambellan, T. Amiranashvili, M. Ehlke, H. Lamecker, S. Lehnert, M. Lirio, N. P. de Olaguer, H. Ramm, M. Sahu, A. Tack, S. Zachow, T. Jiang, X. Ma, C. Angerman, X. Wang, K. Brown, A. Kirszenberg, É. Puybareau, D. Chen, Y. Bai, B. H. Rapazzo, T. Yeah, A. Zhang, S. Xu, F. Hou, Z. He, C. Zeng, Z. Xiangshang, X. Liming, T. J. Netherton, R. P. Mumme, L. E. Court, Z. Huang, C. He, L.-W. Wang, S. H. Ling, L. D. Huỳnh, N. Boutry, R. Jakubicek, J. Chmelik, S. Mulay, M. Sivaprakasam, J. C. Paetzold, S. Shit, I. Ezhov, B. Wiestler, B. Glocker, A. Valentinitsch, M. Rempfler, B. H. Menze, and J. S. Kirschke, "VerSe: A Vertebrae labelling and segmentation benchmark for multi-detector CT images", *Medical Image Analysis*, vol. 73, 2021.

[10] M. T. Löffler, A. Sekuboyina, A. Jacob, A.-L. Grau, A. Scharr, M. El Husseini, M. Kallweit, C. Zimmer, T. Baum, and J. S. Kirschke, "A Vertebral Segmentation Dataset with Fracture Grading", *Radiology: Artificial Intelligence*, vol. 2, no. 4, 2020.

[11] H. Liebl, D. Schinz, A. Sekuboyina, L. Malagutti, M. T. Löffler, A. Bayat, M. El Husseini, G. Tetteh, K. Grau, E. Niederreiter, T. Baum, B. Wiestler, B. Menze, R. Braren, C. Zimmer, and J. S. Kirschke, "A computed tomography vertebral segmentation dataset with anatomical variations and multi-vendor scanner data", *Scientific Data*, vol. 8, no. 1, p. 284, 2021.

Plagiarism Detection in Software Projects Using Abstract Syntax Trees

1st Marek Szymutko

Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xszymu01@vutbr.cz

2nd Pavel Seda

Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
sedap@vut.cz

***Abstract*—Plagiarism is a hot topic in modern education and science. It requires special attention since committing plagiarism is very easy with the use of the internet. This problem can be fought against utilizing prevention or detection methods, which have been both used in this work. This paper introduces an implementation of a submission scheme of students' projects in classes taught at the Brno University of Technology. Scripts for an automatic hand-in space for each student were created. Students have restricted privileges within these spaces on the GitLab cloud service. For plagiarism detection, a tool written in Python was developed. This tool utilizes Abstract Syntax Trees compiled from the source code, which is a part of the Students' solutions. The output of the comparison is represented with a tabular file of the format `.xlsx`, which allows a detailed view. Ongoing implementation is focused on widening the tool's usability by adding a Python similarity comparison engine.**

***Index Terms*—Abstract Syntax Trees, API, Bash, Detection, Git, GitLab, Java, Metrics, Plagiarism, Python**

I. Introduction

This work is related to plagiarism prevention and detection measures used in lectures at the Brno University of Technology. In the subject *Database Systems Security* taught at the Faculty of Electrical Engineering and Communication, students are required to create their own project written in Java (utilizing Maven) or in Python. This application normally has several thousands of lines of code and the number of submissions is around 60 each year.

Given the circumstances, it is impossible to completely prevent plagiarism in the students' work without utilizing automation. The goal of this work is the reduction of uncaught plagiarism attempts.

II. State of the Art

There are many approaches to plagiarism detection of source code and written text. The most versatile (usable both in source code and written text) is metric-based comparison. This technique relies on expressing specific characteristics of compared documents in numbers (e.g., number of lines) [1], [2].

A more advanced universal technique is tokenization. It can also be used both in written text and source code similarity detection. This approach transforms source text or code into different strings which makes the comparison easier. Examples of such algorithms are the compilation of Java code into byte-code, rolling hash, or the WINOWIG algorithm implemented in the MOSS system [3], [4].

Another approach to source code similarity comparison is structural comparison. For this purpose, abstract syntax trees, dependency graphs, or call graphs can be used for comparison. In this scenario, the isomorphism of these structures is searched for. This method eliminates all confusion caused by file refactorization [5]. For this trait, Abstract Syntax Trees were chosen for this work.

Very promising is the behavioral approach to the similarity measure. This compares the program's behavior at runtime rather than the source code itself, which makes it hard to evade [6].

It is also possible to employ machine learning algorithms for similarity detection in source codes. Various kinds of algorithms can be used for this purpose [7].

On the other hand, artificial intelligence can be used for plagiarism evasion also. The rise of ChatGPT shows us that it is impossible to classify AI-generated text as plagiarized because of human-like originality [8].

III. Plagiarism Prevention

The goal of plagiarism prevention measures is to reduce the students' ability to copy each other's code. A project submission scheme was created for this purpose.

At the beginning of the semester, the teacher has to download a table of students that are enrolled in the subject and provide this table to students. All of them have to fill in their GitLab username and the table is then exported to `csv` format. The teacher then prepares the GitLab group for the current year and copies the ID of the group and access token to an `.env` file. They then run a Bash script created by the author of this work and for each of the enrolled students, a separate personal space is created in the specified GitLab group.

These students' spaces are invisible to each other and only the teacher has access to all of them. For the purpose of evaluation, the names of the students' spaces contain the student's name and BUT ID (unique identifier used in the Brno University of Technology information system).

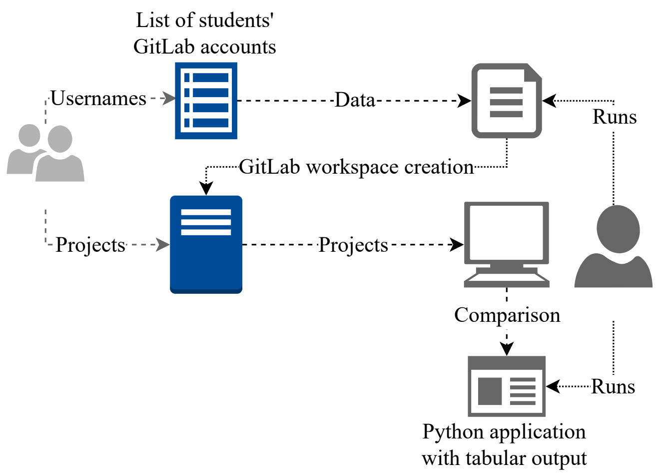


Fig. 1. Simplified Submission Schema

At the end of the semester, students' personal spaces on GitLab are archived in order not to change or delete anything. This is done by degrading the students' role to read-only access via another Bash script.

This way the students are also forced to familiarize themselves with Git versioning system, which is desirable in their future employment.

IV. Plagiarism Detection

This method is focusing on uncovering plagiarized works. In software assignments, it is very easy to copy code and then trying to hide the plagiarized code with small changes. Typical changes done for plagiarism evasion are [9]:

- Change of formatting or comments
- Change of identifiers
- Change of succession of parts of the code
- Change of parts of code to their equivalents
- Structure change of methods
- Combination of own and copied code

The more of these techniques are intercepted by the plagiarism detection algorithm, the better it is.

Plagiarism detection in the scope of this work is performed by a Python application, which communicates with the GitLab API. Because of this, the GitLab group ID and access token have to be provided within a `.env` file for this application also.

The whole simplified submission schema is illustrated by Figure 1.

A. Language Support

As of now, the only programming language supported by the application is Java. But the application is written in modules and an engine for other languages will not be hard to implement. Implementation of the Python similarity detection engine is performed right now.

B. Overview of the Tool

The plagiarism detection tool is run locally. The tool performs the following steps:

- Download of projects from GitLab.
- Search for files with `.java` extension.
- Compilation of files into Abstract Syntax Tree objects.
- Storing the compiled objects in predefined objects.
- Comparison of the predefined objects.
- Presentation of the result.

These steps will be described in detail in the next part of this paper.

C. Download of projects from GitLab

This step also requires credentials for access to the student's projects. The download is performed by `git clone` command with specified credentials and the destination of the clone command. If the download fails (the cloned repository already exists), the command `git pull` is performed in the existing repository instead.

D. Search for files with `.java` extension

Great help for this step is the `pathlib` module for Python. All suitable files are quickly found in the cloned repository.

E. Compilation of files into abstract syntax tree object

For the purpose of AST compilation, module `javalang` was used. This compiler is open-source (licensed under MIT license) and can be downloaded from the portal PyPI.

This module compiles the code like a normal Java compiler, but instead of bytecode, it produces an abstract syntax tree object. This object is then used in the next step.

F. Storing the compiled objects in predefined objects

In this step, AST from the previous step is loaded into a defined data structure. This structure also is in a form of a tree.

The predefined objects sort data types into user-defined and non-user-defined. User-defined types are defined by classes in the submissions, others are tagged as imported or native Java types. But if the identifier of an object was loaded before the definition of the type, the classification would fail and the object would be falsely labeled as non-user-defined. To solve this issue, only the string of the identifier name is stored and the loading of the object is done via *cached properties*. In Python, properties are similar to *getters*. Cached properties are initialized once only and then they store the value. This way the full loading of the object only appears during the first comparison, which makes the algorithm reliable and fast.

All objects in the predefined structure have a method `compare` which is used to determine the similarity of the objects. This way similarity is computed hierarchically from the root (object representing the student's project) to its leaves (for example data types, methods…).

G. Comparison of the predefined objects

The predefined tree structure is compared via a method called `compare`. This method implements pairwise comparison which is different for every type of node in the tree. For example, nodes that represent methods of the Java code are compared based on their interface (visibility, arguments, and return type) and statements (parts of the code divided by semicolons).

The exact implementation for each object is out of the scope of this paper and can be found in the source code of the application [10]. But in no object, the name is used for comparison. That is because, in modern development environments, it is very easy to refactor the project in an attempt to mask plagiarism (as discussed in section IV).

Calls inside method bodies within compared projects are also inspected to check if the called method is in the scope of the project and if yes, statements of this called method are also counted towards the parent method (this way code relocation will not mask plagiarism well enough).

The return type of the `compare` method is a type called `Report`. This object also has a tree structure, but rather than mapping projects to a tree structure, it maps similarities between pairs of projects. Each report has a **similarity score** (integer between zero and one hundred), a **weight** (positive integer based on the size of the compared tree), **the two objects, which are compared**, and **child reports** (in order to hold the tree structure).

The algorithm for finding the best match of objects between two projects can be described as follows:

1) All child nodes of the compared object are matched to other objects' child objects via the Round-robin algorithm.
2) The best match is selected based on the similarity score. All other matches containing either of these child objects are discarded. This way every object is matched once only.
3) The algorithm is repeated until it runs out of uncounted child objects. If the number of the child objects is not the same between the two parent objects, the remaining objects will be matched with `NotFound` object and graded with a score of zero.

This algorithm can also be illustrated with Table I (in this example, methods of a pair of classes are compared and 2 iterations are performed).

H. Presentation of the result

This part is about the presentation of the result of the application run.

The option chosen for the outcome representation was a tabular file of the `.xlsx` format. The first sheet in the application contains an overall table where every project is compared to every other project and graded with the similarity score. There also is a review of the GitLab groups where no project was found and an assessment that does not contain any `.java` files.

TABLE I
MATCH-FINDING ALGORITHM

Methods	`createBook`	`deleteBook`
`updateCar`	~~15~~	42
`deleteCar`	~~14~~	87
`createCar`	**97**	~~35~~

↓

Methods	`deleteBook`
`updateCar`	~~42~~
`deleteCar`	**87**

	student A	student B	student C	student D	student E	student F	student G	student H	student I	student J	student K	student L	student M	student N	student O
student A		75	52	73	45	64	45	73	53	62	44	75	67	67	75
student B	75		55	71	47	64	37	65	51	54	44	65	72	66	92
student C	52	55		53	27	42	37	51	46	38	37	50	51	52	55
student D	73	71	53		43	69	47	79	60	61	39	79	78	78	71
student E	45	47	27	43		51	22	43	34	45	40	43	47	37	48
student F	64	64	42	69	51		43	67	53	67	46	63	66	64	60
student G	45	37	37	47	22	43		53	44	33	34	46	42	48	36
student H	73	65	51	79	43	67	53		59	57	41	76	69	76	65
student I	53	51	46	60	34	53	44	59		44	47	57	55	56	51
student J	62	54	38	61	45	67	33	57	44		40	67	60	50	50
student K	44	44	37	39	40	46	34	41	47	40		40	39	43	43
student L	75	65	50	79	43	63	46	76	57	67	40		76	68	60
student M	67	72	51	78	47	66	42	69	55	60	39	76		68	71
student N	67	66	52	78	37	64	48	76	56	50	43	68	68		65
student O	75	92	55	71	48	60	36	65	51	50	43	60	71	65	

Fig. 2. Similarity Representation Overview

The table on this sheet is color-coded according to the score. Due to the nature of the submissions (similar tasks), the score can be usually pretty high even though it does not show any signs of copy-pasting of the code. For this reason, the chosen threshold and colors were:

- Up to 70/100: green
- Up to 85/100: yellow
- Other: red

Each of the score cells works as a link to a more detailed view, which is on a different sheet of the file. Every detailed view (every pair of projects) is on a separate sheet.

Part of the overview sheet can be found in Figure 2 and part of the detailed view sheet is depicted in Figure 3.

V. CASE STUDY

During the academic year 2022/23, there were 60 students enrolled in the Database Systems Security subject. 42 of these students submitted an assignment written in Java before the deadline of the projects. The project was not submitted by

JavaFile	LoginView.java	PersonAuthView.java			95
	JavaClass	LoginView	PersonAuthView		95
		JavaMethod	setUsername	setEmail	100
		JavaMethod	setHashed_password	setPassword	100
		JavaMethod	getUsername	getEmail	100
		JavaMethod	getHashed_password	getPassword	100
		JavaMethod	NOT FOUND	toString	0
JavaFile	AppEditController.j	BookEditController.java			89
	JavaClass	AppEditController	BookEditController		89
		JavaMethod	handleEditPersonButton	handleEditBookButton	100
		JavaMethod	personEditedConfirmationDialog	personEditedConfirmationDialog	100
		JavaMethod	loadPersonsData	loadPersonsData	100
		JavaMethod	setStage	setStage	100
		JavaMethod	initialize	initialize	72

Fig. 3. Similarity Representation Detail

15 of the students. One person chose to write the assessment in Python and two students submitted just concepts for their assignment.

The application was run on a laptop with CPU Intel i5 8300-H CoffeeLake. The Run took approximately 12 minutes. During this time, all projects were cloned from their respective GitLab repositories, parsed, and compared and the output was stored in a `.xlsx` file of size approximately five megabytes.

Out of 42 students, 13 had no suitable match found (all their cells in the tabular sheet were marked as green). On the other hand, six students were marked by red color on at least one match. The other students most likely copied some idea either from each other or from the template project provided by the teacher of this subject but did not copy-paste most of the project.

Most of the time it is impossible to undoubtedly determine which assignments were plagiarized. The automated means searching for code similarities, which could also occur randomly. Even if some work is most likely plagiarized, it is hard to decide, which student copied from whom. One lead could be timestamps of the commits which introduced similar code, but these could be manipulated. It is up to the teacher to handle the results of the similarity test according to their best judgment.

The highest similarity score from all 42 students was **95/100**. In this match, most of the files were showing signs of copying (the majority of files between the two projects were the same, only with refactored identifiers of the variables used). These files were also the same as files within the example project provided by the teacher of the course. Copying parts of the example project was not forbidden but discouraged.

Out of the six students marked with red color, four handed assignments similar to the example project. The other two students submitted projects that were very similar to each other (score 92/100) but not to any other project (the second highest match was 75 for both of these students). This shows that these students most likely shared the code in order to help each other. When confronted with this result by the teacher, these students confessed to plagiarism.

Most of the students, whose submissions were marked with yellow color at most, uploaded projects that had some characteristics similar to the example project provided by the teacher. This is not against the rules of the course, because students are not required to write all parts of the projects from scratch.

One student also uploaded a solution containing a syntax error, which made the comparison application skip this invalid file. One other student used syntax not compatible with the AST compiler in several files, which made the application skip these files. The reason is that the compiler `javalang` is written for Java 8 and students were allowed to use a newer version of Java for their projects. This is unfortunate and most likely will not be fixed unless a better AST compiler for Java is released. This limitation only affects files that use specific syntax from newer versions of Java. Files that cause compiler errors are logged and should be checked manually afterwards.

ACKNOWLEDGMENT

The work described in this paper was done for the purpose of a bachelor's thesis. The supervisor of the thesis is Pavel Šeda whom I thank for his professional lead and sincere attitude.

REFERENCES

[1] Romans Lukashenko, Vita Graudina, and Janis Grundspenkis. 2007. "Computer-based plagiarism detection methods and tools: an overview," *In Proceedings of the 2007 international conference on Computer systems and technologies (CompSysTech '07)*. Association for Computing Machinery, New York, NY, USA, Article 40, 1–6. doi: 10.1145/1330598.1330642

[2] Thomas Lancaster & Fintan Culwin. 2005. "Classifications of plagiarism detection engines", *Innovation in Teaching and Learning in Information and Computer Sciences*, 4:2, 1-16, doi: 10.11120/ital.2005.04020006

[3] J. -H. Ji, G. Woo and H. -G. Cho. 2008. "A Plagiarism Detection Technique for Java Program Using Bytecode Analysis," *2008 Third International Conference on Convergence and Hybrid Information Technology*, Busan, Korea (South), pp. 1092-1098, doi: 10.1109/ICCIT.2008.267.

[4] Alex Aiken. 1985. "Moss: A System for Detecting Software Similarity," *Stanford University*, Stanford, California

[5] Hyun-Je Song, Seong-Bae Park, Se Young Park. 2015. "Computation of Program Source Code Similarity by Composition of Parse Tree and Call Graph," *Mathematical Problems in Engineering*, vol. 2015, Article ID 429807, 12 pages, doi: 10.1155/2015/429807

[6] H. Cheers, Y. Lin and S. P. Smith. 2021. "Academic Source Code Plagiarism Detection by Measuring Program Behavioral Similarity," in *IEEE Access, vol. 9*, pp. 50391-50412, doi: 10.1109/ACCESS.2021.3069367.

[7] A. Eppa and A. Murali. 2022. "Source Code Plagiarism Detection: A Machine Intelligence Approach," *2022 IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, Bengaluru, India, pp. 1-7, doi: 10.1109/ICAECC54045.2022.9716671.

[8] M. Khalil and E. Er. 2023. "Will ChatGPT get you caught? Rethinking of Plagiarism Detection" *ArXiv*, Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway and Middle East Technical University, Turkey, doi: 10.48550/arXiv.2302.04335.

[9] G. Whale,. 1990. "Identification of Program Similarity in Large Populations", *The Computer Journal, Volume 33, Issue 2*, Pages 140–146, doi: 10.1093/comjnl/33.2.140

[10] Marek Szymutko. 2023. "Plagiarism detection tool for FEEC BUT," *GitLab* https://gitlab.com/but-courses/bpc-bds/instructors-supporting-tools/bds-similarity-check.git

Back to the Future: Developing an On-Board System for a Vintage Motorcycle

Jan Sevcik*, Radovan Juran†
Brno University of Technology
Faculty of Electrical Engineering and Communication
Brno, Czech Republic
*xsevci71@vutbr.cz, †radovan.juran@vut.cz

***Abstract*—This paper describes complex design of a prototype of an on-board system for older motorcycle for measuring its operational characteristics. The birth of the idea is described, followed by explanation of characteristics measured by the system, principles of their measurement, and processing the data for visualization on a display. The system is designed with the aim of the most minimal intervention into the construction of the motorcycle, but being easy to reproduce and durable enough for a real deployment and use while riding the motorcycle.**

***Index Terms*—on-board system, motorcycle measures, flowmeter, esp32**

I. Introduction

Vintage motorcycles have a unique appeal, but their traditional designs usually lack modern features and technologies, especially measuring and monitoring operational parameters, such as fuel consumption. This work aims to address this issue by developing an on-board system for an old motorcycle that enhances its functionality, proposing a system to bridge the gap between classic motorcycle designs and modern standards, providing useful data of the motorcycle operation statistics, long-term analysis of costs and planning the needed maintenance.

The paper describes the design of on-board system, its benefits, and the challenges encountered during the process of a complete design of on-board system for motorcycle from 1960s. The very first idea for this project stemmed from need of adding some kind of measuring, because the target motorcycle lacks any kind of measurement, besides the mechanical speedometer and tachometer. Then came an idea to improve power of the original engine by enhancing or replacing some key parts of the engine for overall increase in the engine power. In this case, the system can be utilized also for a performance comparison by using the measured data from the original engine characteristics and after its performance fine-tuning.

The on-board system is designed with minimum intervention into the construction of the motorcycle in mind, so that anyone can re-use this idea, improve it and adapt for any other similar type of a motorcycle. The on-board system has to provide easy to use operation, so that any user can easily find all measured characteristics.

II. Proposed design

The system is meant for use in a specific motorcycle – a 1969 Jawa 250/592 [1]. The whole central unit of the system including the display is designed to fit in the already existing hole on top of the gas tank, where the switch box is placed from the factory. However, as most of Jawa motorcycles have same or similar construction, it's possible to mount all parts of system on many other motorcycles of similar type. The only difference is the location of the central unit. There are only a few models sharing the same shape of body parts, so it might be needed to redesign the housing for the central unit and mount it on the handlebars or somewhere convenient.

From the central unit cables will lead to the sensors and actuators, as seen in the Fig. 1, intended to measure the following data:

- speed,
- engine rotation speed (rpm),
- fuel consumption,
- temperature of cylinder head and environment,

and also enable following features:

- replace the function of the switch box (starting the engine, turn on/off lights),
- transfer acquired data via Wi-Fi (e.g. to a mobile phone).

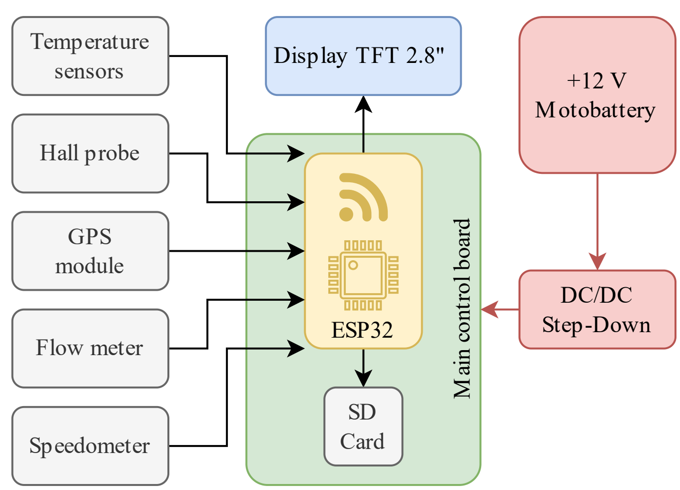


Fig. 1: Block schematic diagram of the onboard system.

DOI 10.13164/eeict.2023.49

The Fig. 2 is showing real intended placement for system housing in gas tank switchbox hole.

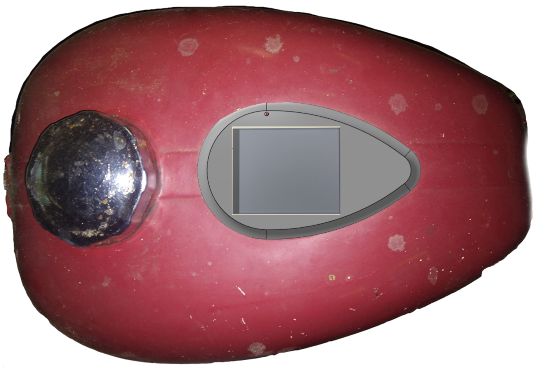


Fig. 2: The intended placement in the gas tank.

A. Central element – ESP32

After a market research the ESP32 by Espressif Systems was chosen to become the core of the system, especially for its price to performance ratio. The well known ESP32. It belongs to a series of cost-effective, low-power SoC (system-on-a-chip) with Wi-Fi and Bluetooth connectivity, released in 2016 and still in production with new revisions. Its predecessor ESP8266 had some bugs and security problems, which ESP32 managed to solve and bring better performance [3]. Development kits are available as well as just the modules. Although there are other commercially available options, the proposed system aims on maximal cost rffectivity and replicability of the whole system by hobbyists and easy upgradability thanks to large ESP platform community. Because the available off-the-shelf boards are intended primarily as development kits, they have poor connection options for application in the designed motorcycle system, therefore a custom board will be designed to secure better ways to connect the planned sensors in a mechanically durable way.

B. Display

Many modern motorcycles utilize TFT displays, therefore a quite similar size and type was chosen – a 2.8 inch display, also thanks to its low price and decent quality [4]. Also thanks to chosen TFT display there is big possibility of maximazing high level graphical output. A slight disadvantage lies in terms of observability, which could be bad especially on the sunlight. After field testing there its still possible that visibility will be too bad and change of display will be needed. Possible solution to this issue could be an OLED display, which offers better observability, but these come only in smaller sizes and visibility of measured quantities would be bad from a greater distance when riding motorcycle.

C. Flow meter

In this case the market research showed that there are not many available products meeting the needs to measure fuel flow on the Czech market available in retail sale, thus it was necessary to search abroad and contact the manufacturers. One of the best suitable flow meters found is the FCH-m-POM by a German company named B.I.O-TECH e.K. [5]. This company produces a large portfolio of flow meters and controllers so it was possible to find several products which could be suitable for the target application. After consultation with the company sales team the best solution was found, which also provides good enough IP coverage (IP 65), as the flow meter is located on an exposed place where it can easily get damaged by water. The flow meter FCH-m-POM offers discernment from 0,015 to 1 l/min and is suitable for use for water, diesel, oil and gasoline for price about 44 EUR. The measurement principle is based on Hall-effect [6]. For connection only three pins are required: GND, VCC power pin and Signal (data) pin.

D. Tachometer

System for measuring the rotation speed of the engine is based on measuring impulses induced on induction coil. The custom designed board is based on an existing stroboscope circuit design [7], modified to measure the motorcycle engine ignition timing (see Fig. 3).

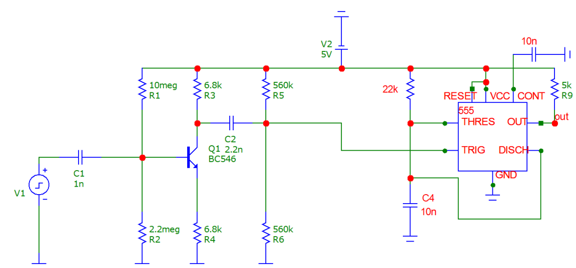


Fig. 3: Circuit for tachometer.

Signals are brought to the one-stage amplifier in connection with a common emitter. C1 separates the DC component and signals are linked into flip-flop input based on NE555 timer. Simulation of the circuit can be seen in the Fig. 4.

First prototype of the tachometer was made by THT components and tested in laboratory conditions by 1 kHz input from Agilent 3320a signal generator and Agilent DSO1012A oscilloscope. The measurement showed that design of the board works as intended and its possible to move on next step which is field testing on real motorcycle by winding the coil around the high-voltage cable of the motorcycle and connecting it to the oscilloscope. By laboratory measures with oscilloscope and waveform generator was managed to test that minimum voltage needed for correct functionality is only about 0.5 V when powered by 5 V DC. Oscilloscope measuring corresponds with simulated waveforms in the Fig. 4, as can be seen in Fig. 5. Because the prototype worked correctly, it is possible to optimize its size using SMD components and perform testing of its functionality on the motorcycle.

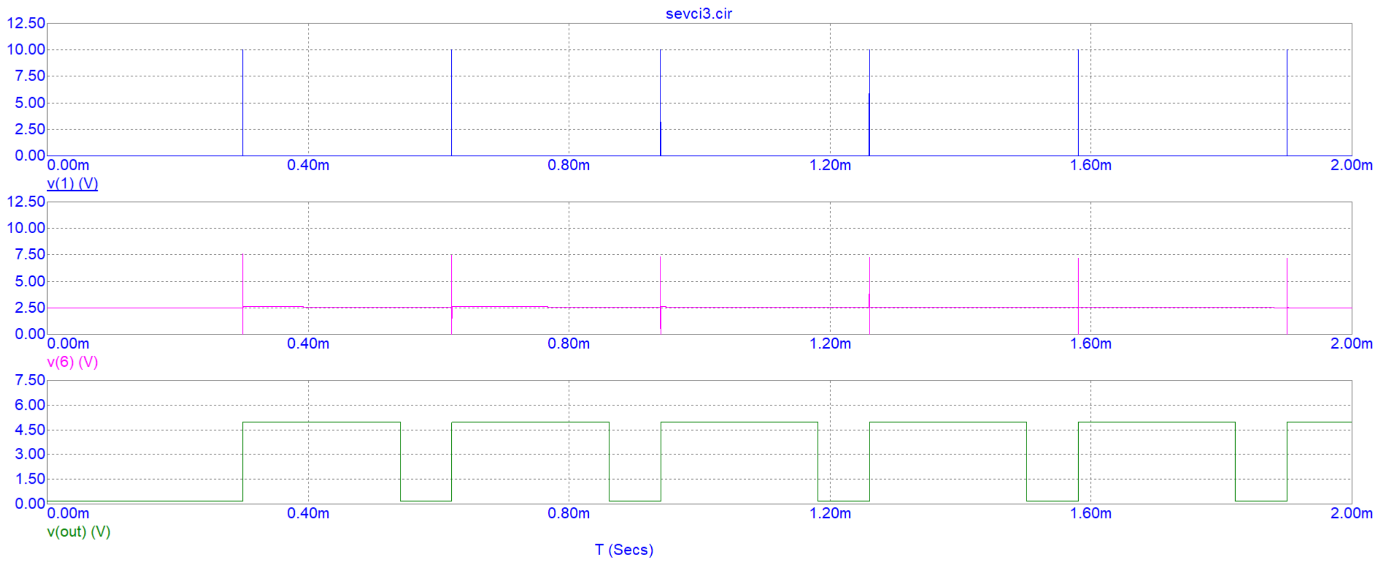


Fig. 4: Circuit simulation for tachometer in Microcap tool.

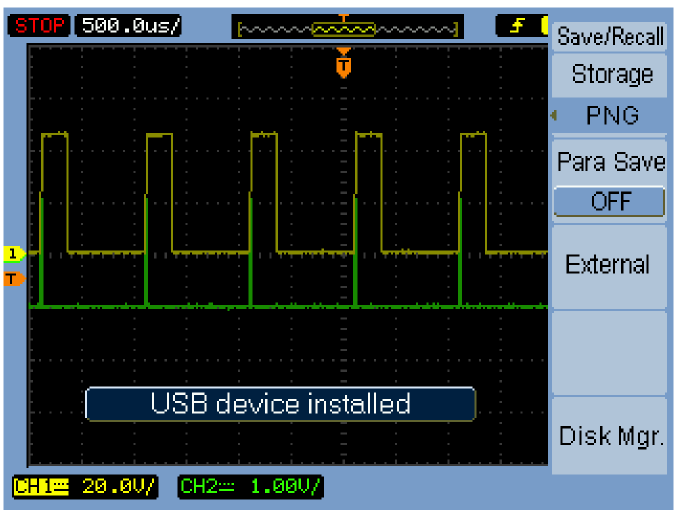


Fig. 5: Laboratory measuring of the prototype board.

E. Storage

The measured data are primarily intended to be shown on display, but also have secondary use for a long-term analysis. To store the measurements a micro SD card in SPI mode is used. All files should be accessible by either removing micro SD card and reading in an external device or directly from the system via web server run on the ESP32.

F. GPS module

GPS coordinates are vital part of measured data, utilized for safety (anti-theft) and trajectory history for driving record log book entries, which can be afterwards transformed on a map route. Hardware implementation takes into account the GYNEO6MV2, a cost-effective GPS module with included antenna [8].

G. Speed and distance

For measuring motorcycle speed and travelled distance a Hall sensor was chosen to use, intended to be mounted on the suspension of motorcycle and magnet near located on the motorcycle wheel. It opens a way to measure data other way than the old mechanical method made by factory. It uses magnetic field sensor NJK-5002A [9], but it needs to be further tested, especially to find out whether the sensor can match the motorcycle speed. It also provides IP68 which ensures it will survive in its location.

H. Temperature

On the motorcycle there are two temperature sensors measuring the temperature of the cylinder head and the environment. Sensor for measuring environmental temperature is BMP280 [10], which also provides atmospheric pressure, which can be used to calculate altitude. Sensor will be located in the body of the motorcycle, protected against water and ensuring not to be affected by air flow. The second sensor is located on top of the cylinder head near spark plug. Because engine is two stroke, its not needed to measure temperature of oil, but it is preferred to measure temperature of cylinder and its head, because its air cooled engine.

Because its chosed undestructive way its ensured measurement will be affected and temperature will be lower than real. It can be fixed by measuring real temperature and improving code on ESP32. Because of the high temperature of head we need to choose better temperature sensor which allows measure higher temperatures. Its chosen to use KTY81 [11] which measures up to 150 degree Celsius.

I. RFID

To ensure that motorcycle can be started and controlled only by its owner or allowed persons it was decided to add an RFID module RC522 using SPI interface.

III. Control system

Whole on-board system needs to have low power consumption while motorcycle is not used. That is achieved by switching ESP32 to its deep sleep mode. During this sleep mode some pins of ESP can be controlled by ULP co-processor and thanks to that power consumption is about 100 μA [12]. During sleep mode some of pins can be used, namely RTC GPIO pin. There are several options to wake ESP up but we will use one of the pins combined with a button. After waking up ESP prompts the user for authentication via RFID module. After successful authentication ESP will fully turn on and unlock additional control options.

First option will be by switching on the power through a low level trigger optocoupler relay module to allow user start the motorcycle. During this time all modules will start to measure and collect data which will be saved on micro SD card. Second option will be allowing to turn lights on, which will be also provided by relay module and can be anytime turned off by button located next to main display.

Also, an optimal solution for easiest and convenient data output has to be taken into account. Besides display, the current solution considers the use of a web-server running on the ESP and being accessible by a hot-spot via internal Wi-Fi module, for example by using a button for limiting time that ESP uses for scanning nearby networks. On the web server a basic GUI (graphic user interface) is considered with data overviews and ability to manage the saved data on the micro SD card.

IV. Designing PCBs

This work contains of parts which need its own specific custom board design. All board are designed in open source EDA (Electronic Design Automation) software suite named KiCad [13], currently in version 7.0.1.

The first board is tachometer which is more described in section II-D. This board will be reduced in size by using SMD components and add together with main control board with ESP32 in the next version of the szstem.

To ensure the most optimal and reliable connection of all the modules and sensors to the main board, the right selection of connectors needs to be considered. Two options taken in account are currently either using screw terminals which will ensure strong connection that lasts vibration produced by motorcycle and provide easy assembly.

t also need to be keep on mind that casing of the modules need to provide enough protection against water and moisture not only from above where display should be located but also a grommet hole for cables from modules and sensors.

V. Conclusion

The paper introduced a complex design of a on-board system for an old motorcycle, which aims to keep the historical looks nearly intact, yet bring in new features, ensuring a non-destructive mounting on the motorcycle body. The individual sub-systems are described, some of which were designed from scratch including simulation of their operation. Additional work focuses also on renovation and repairs of motorcycle mechanical and body parts. After completing motorcycle it will be possible to test all sensor and whole system on motorcycle and make necessary improvements. Future work will focus on ensuring that the system provides durability and reliable functionality in real scenarios. By choosing ESP32 with software development based on Arduino framework it is possible to easily add a many other modules and their combinations to improve the system in the future.

References

[1] *Jawa 250/592* Wikipedia [online]. [cit. 11. 03. 2023]. Online: ¡https://cs.wikipedia.org/wiki/Jawa_250/592¿. II

[2] *ESP32* Espressif [online]. [cit. 11. 03. 2023]. Online: ¡https://www.espressif.com/en/products/socs/esp32¿.

[3] *ESP32 vs ESP8266 – Pros and Cons* Makeradvisor [online]. [cit. 11. 03. 2023].Online: ¡https://makeradvisor.com/esp32-vs-esp8266/¿. II-A

[4] *TFT Display* [online]. [cit. 11. 03. 2023]. Online: ¡https://www.aliexpress.com/item/1005003797803015.html¿ II-B

[5] 97478084-FCH-m-Pom *BIO-TECH Flowmeters* [online]. [cit. 11. 03. 2023]. Online: ¡https://www.btflowmeter.com/fileadmin/PDF/Flowmeter/97478084_FCH-m-Pom.pdf¿ II-C

[6] *Princip Hallova senzoru* [online]. 2015 [cit. 11. 03. 2023]. Online: ¡https://e-konstrukter.cz/novinka/princip-hallova-senzoru¿. II-C

[7] Tomáš Brožek *Jawa50.cz Stroboskop s LED (pro seřízení předstihu)* [online]. 2018 [cit. 11. 03. 2023]. Online: ¡https://www.jawa-50.cz/clanek/navod-na-vyrobu-stroboskopu-serizeni-predstihu-led-sonda.html¿ II-D

[8] *NEO-6 series* [online]. [cit. 11. 03. 2023]. Online: ¡https://www.u-blox.com/en/product/neo-6-series¿. II-F

[9] *NJK-5002C Hall Effect Proximity Sensor* [online]. [cit. 11. 03. 2023]. Online: ¡https://www.handsontec.com/dataspecs/sensor/NJK-5002C-Hall%20Sensor.pdf¿. II-G

[10] *NEO-6 series* [online]. [cit. 11. 03. 2023]. Online: ¡https://www.bosch-sensortec.com/products/environmental-sensors/pressure-sensors/bmp280/¿. II-H

[11] *KTY81 series* [online]. [cit. 11. 03. 2023]. Online: ¡https://www.nxp.com/docs/en/data-sheet/KTY81_SER.pdf¿. II-H

[12] *ESP8266 Deep Sleep with Arduino IDE* [online]. [cit. 31. 03. 2023]. Online: ¡https://randomnerdtutorials.com/esp8266-deep-sleep-with-arduino-ide/¿. III

[13] *KiCad EDA A Cross Platform and Open Source Electronics Design Automation Suite* [online]. [cit. 31. 03. 2023]. Online: ¡https://www.kicad.org/¿. IV

Rule-based Engine for Position Data: Low-code/No-code Approach to Process Logic

Natalie Vychodilova*, Radovan Juran †
Department of Telecommunications
Brno University of Technology
Brno, Czech republic
*230911@vutbr.cz, †radovan.juran@vut.cz

***Abstract*—Real-time location systems (RTLS) are becoming a standard part of modern digitized industrial facilities, enabling acquistition of position data of objects or people tracked in a given monitored area. These systems can utilize the virtualized coordinates to divide the real space into user-defined digital zones and give them meaning within a process. This paper describes the design and implementation of a commercial RTLS system extension, intended to eliminate the need for programming any complex position data processing by leveraging Low-code/No-code approach to creating process logic in a user friendly way.**

***Index Terms*—Low-code / No-code rule-based engine, RTLS, positioning systems, safety, efficiency, production process**

I. Introduction

Localization systems in general use an infrastructure of a given indoor positioning technology to monitor the movement of objects in confined spaces and execute logical processes over the gathered data, providing valuable feedback to boost the efficiency and safety. These systems enable companies to control a wide range of critical activities such as demand management, sales planning or tracking production and logistics processes [1]. However, due to the amount of data that is collected, these systems are becoming very complex and sophisticated. As a result, the end user finds these systems difficult to use and requires specialist knowledge to operate them. Plant managers, responsible for ensuring smooth run of production process, often lack programming skills, yet they understand the production and logistics processes and define correctly their needs and expectations.

The solution to help these workers is to create a user-friendly interface, where the plant manager can create custom rules over the collected and processed position data to control specific manufacturing processes without the need for knowledge of system logic or programming.

II. Process problems

Often there are certain risks associated with any manufacturing or logistics process. Although these risks can never be completely eliminated, they can be predicted and prevented. A process problem is an unexpected event that is a threat to the safety of people or the process itself. This event is unpredictable and is usually preceeded by a sequence of unexpected events, which may be caused by a machine, human factor, or an interaction of both.

A. Problems in manufacturing industry

Any problem that occurs is undesirable and has a major impact on the operation of the factory. Whatever the cause of the process problem, it will somehow slow down the production process, stop it completely or cause damage. The human factor causes can be events such as falls, collisions, or use of damaged tools. These events can result in an injury to the operator or damage to the machine.

When multiple machines collide, or a machine and an operator collide, both the safety of the operator and the operation of the production process are compromised. These problems arise in shop floors as well as in warehouses, altough modern warehouses are built to maximally remove the human factor from the workplace. Most tasks are carried out by machines programmed with clearly defined steps, resulting in an optimiziedefficient logistics processes. The operator only monitors the process and intervenes only when an unexpected event occurs.

The occurrence of a process problem is not always attributed to collisions between machines or ongoing operations. We can also look at the problem from a different perspective: Consider an office building and the people working there. Not all of them have access to all areas of the building. If an employee moves to an area without permission, that can also be taken as a type of a process problem.

B. Solution

It is important to remember that every problem that occurs costs time and money. Any process problem slows down or even stops the production flow and the time it takes to return to the pre-event state is undesirably costly. The question is how to prevent these events, ensure the continuity of the processes and control each individual object and process.

Today, advanced positioning systems are used in production and logistics and are becoming an integral part of a company's critical infrastructure [2]. By adding a rule-based engine to an existing RTLS system, it is possible to ensure the continuous monitoring of the correct execution of all the rules that control the individual processes.

DOI 10.13164/eeict.2023.53

III. Rule-based engine requirements

The proposed system is an extension to comertially deployed RTLS system by Sewio and leverages its capabilities of a positioning platform. The system server collects data on the position of individual objects moving in the an enclosed space. Based on the time stamps, it can also determine the speed of the object and its trajectory. This data can be stored and used for later analysis, for example to optimize travelled trajectories in the facility.

The tracked objects can be divided into four categories: people (persons), logistics elements, materials (products) and tools (machines). The division comes from the production processes and gives a better overview of the monitored scenario. In the proposed rule-based engine, specific variables are defined for each tracked object type. Examples of variables are shown in the Tab. I.

TABLE I
Examples of variables for each object type

Object type	Examples of variables
people	*type of person* (employee, manager, visitor), *working status* (working, break, lunch break)
logistics elements	*type of element* (trolley, automatic guided vehicle), *operating condition*, *battery level*
materials (products)	*schedule, workflow, total workflow time*, *factory space, product temperature*
tools (machines)	*operational status* (in use, available), *permissions* (personnel who can use the product)

a) Types of objects: Different rules are required for each type of object listed in Tab. I. For example, for an object of type *materials (products)*, user may want to create a rule that checks that the product has not repeated a step in the manufacturing process. However, this rule may not be useful for an object of type *people (persons)*, because a person does not go through a production process, but only manages it.

b) Logic: A rule-based engine will need to translate simple (verbally interpretable) information entered by the user into computer logic. If the interpretation is more complex, the logic will be more difficult to develop. It is possible that the user may express the rule incorrectly, or that the condition is too long and will take longer to check. These problems need to be anticipated and avoided. The solution to the complexity of these rules is to create templates that users can use, rather than forcing them to create their own rules. The use of defined templates also avoids the creation of duplicate rules. The administration of the rules themselves is simplified.

c) Time: Time adds complexity to the system (or the part of the system that will evaluate the rules being created). The system will combine asynchronous data streams (real-time received data) with synchronous information (end-to-end query API). The problem with the time is that the condition may not be current at the time the rule is executed.

Imagine a production hall where a machine is placed and a zone is created around it (see Fig. 1). A rule is created for this zone: *If a person enters the zone, the machine will slow down and an alarm will sound*. Suppose a person does not hear this

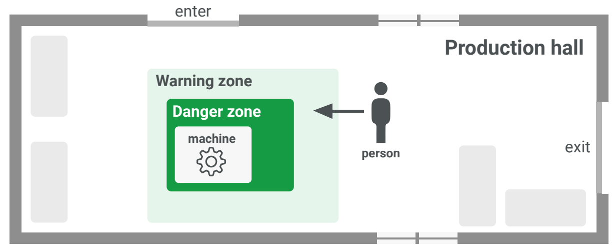


Fig. 1. Situation in the production hall.

warning and comes dangerously close to the machine. In this case, the machine must stop completely. This means that the proposed system must predict the future to some extent while working with the past. *(Is the condition still valid? How long will the condition be fulfilled?)*

d) Uncertainty: Uncertainty arises from this scenario. Suppose the person in the previous case did not hear the warning message and continued to approach the machine. However, there is some probability that the person will react to the warning message and move away from the machine. The proposed system should operate with this probability. The uncertainty is also affected by the actual transmission of the signal, as it is transmitted wirelessly from the tags. The life of these hardware devices is dependent on the battery.

IV. Structure of rules

Suppose the user wants to create a rule based on the situation shown in Fig. 1. Scenario: The user wants the self-service machine to slow down if a person enters the *danger zone* and, at the same time, if the operator enters the *warning zone*, the machine must stop completely and the warning alarm must sound.

The question is how to go about creating rules for this situation. One possibility is the translation of a verbally interpreted situation into a single condition, but this is not a suitable in terms of speed of rule evaluation. It is more efficient to split the situation into two, as follows:

- The user wants to slow down the machine if a person enters the warning zone.
- The user wants to stop the machine completely if the person enters the danger zone.

```
Scenario: The user wants to slow down the machine
if a person enters the warning zone.

given:  status of the zone (empty, occupied)
when:   zone is occupied
then:   reduce the speed of the machine
```

Listing 1. Structure of a scenario-based rule.

Splitting a complex rule into two simpler ones reduces the amount of time for the engine to evaluate the rule, makes it easier to filter out individual rules, and prevents the duplicity of the same or mutually exclusive rules. List. 1 shows that the rule consists of two parts. The first part is the definition of a condition (when), the second part of the rule (then) is executed only if the conditions defined by the user are met.

A. Trigger

We will refer to the defined rule condition as a trigger. A trigger is an event that activates the rule. The event can be periodic (the rule is triggered periodically at a specific time), manual (the rule is triggered manually by the user administering the system) or automatic (the rule is triggered by a trigger). A rule can contain several triggers. These triggers are separated by logical AND, OR or NOT operations.

Examples of triggers that activate rules:

- Product has been in the production process for too long.
- Tag battery level is less than 20 %.
- Product has skipped a production step.
- Person is close to the searched tool.

Our proposed rule-based engine regularly queries the existing real time locating system system. Sewio's RTLS Studio supports interfacing via REST and Web Sockets APIs and provides information about zones, buildings and moving objects. By connecting the rule-based engine to RTLS Studio via the Web Sockets API, we get the data needed to create a rule. These variables are passed to the system, which evaluates the trigger value based on the condition.

B. Action

The output of the rule is an action. It is executed on all objects that meet the user-defined triggers. By action we mean an event that can intervene and influence the operation of a production or logistics process, alert operators and workers, manage administration or schedule events. Just as a trigger can trigger several actions, a single action can be triggered by several triggers.

Examples of actions performed after the rule is activated:

- Schedule a battery change in the calendar.
- Slow down a moving machine.
- Send a notification to everyone in the zone.

V. Model of Rule-Based Engine

The system processes rules created by user in the web interface. The system is divided into parts, each with its own function. All parts of the system communicate with each other and data is transferred through their interfaces. Fig. 2 shows the model of the system.

The core is a rule-based engine that evaluates user-created rules. A web-based user interface that communicates with the core system is used to create, manage and visualise rules. API connectors (REST and Web Sockets) connect the rule-based engine to the Sewio RTLS software. Using these connectors, the system queries and evaluates the responses received from the RTLS system. All rules created, including pre-created templates and user data, are stored in a database that communicates directly with the system.

A. Database

The database stores information about users and the history of changes made to the rules. Storing the history of rule changes and executions is important for later analysis of the rule-based engine. By analysing the activation of individual rules, production and logistics processes can be optimised. Instead of storing data in a relational database, it is preferable to store the data of the proposed system in a non-relational document-based database. An example of such a database is MongoDB. The reason why it is more efficient, and to some extent necessary, to store data in documents is that not all data in our system has the same structure. In our case, we are talking about *trigger* and *action* objects - each of these objects will have a different structure.

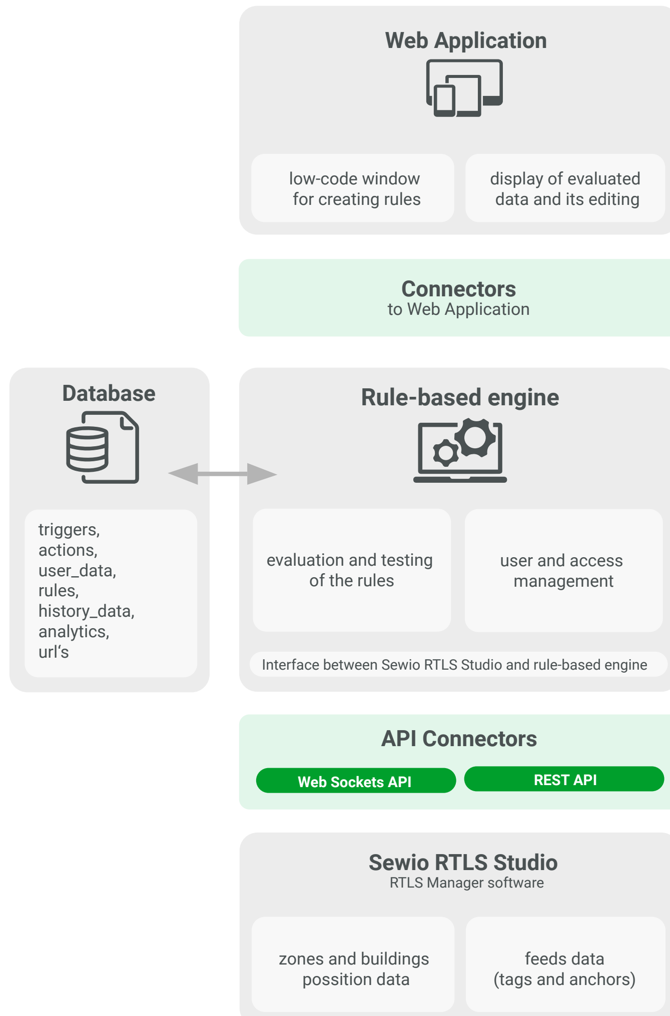


Fig. 2. Model of the rule-based engine for position data.

List. 2 below shows the storage of two triggers. The first triggers a rule when an object enters the zone. The second trigger activates the rule if the object stays in the zone for more than a certain amount of time. Each trigger has a different set of information requirements for activation. The *object_enters_zone* trigger only needs the zone id. The *object_in_zone_longer_than* trigger requires a variable to hold the time in addition to the zone id.

```
{
  "object_enters_zone": {
    "id": "1",
    "description": "When object enters zone",
    "zone_id" : "42",
    "function" : "zone_status(zone_id)"
  }
  "object_in_zone_longer_than": {
    "id": "2",
    "description": "When object is in zone longer
    than time",
    "zone_id" : "16",
    "time_in_zone": "15 sec",
    "function" : "in_zone(zone_id, time_in_zone)"
  }
}
```

Listing 2. Structure of stored triggers.

B. User interface

The user interface allows the user to create rules on position data. These rules can be managed (edited, deleted, filtered, searched) by the user. The user should also be able to test the rules that have been created via the web interface. The user interface is simple, clear and responsive. Different user groups have different permissions. The administrator is allowed to manage individual users and perform operations with rules.

After logging in to the web application, the user is redirected to a dashboard screen (Fig. 3).

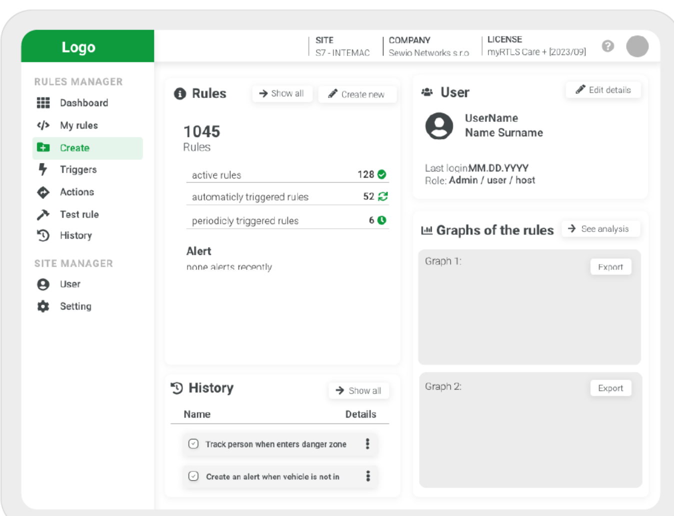


Fig. 3. User interface – dashboard.

It displays all important information such as a table containing a list of active rules, a history of recently executed rules, and graphs showing the analysis of rules. The user can create rules from scratch or by using the preset templates that are stored in the system. On the left side of the page, a menu is displayed that links to the different pages of the system.

The web UI communicates with the rules-based engine using service components. The communication works on the principle of querying and displaying answers in the web environment. The web interface is built on a low-code platform that facilitates the development and editing of individual interface elements [3]. The rule creation window is also based on the low-code principle [4], [5]. The end user drags the individual blocks into a continuous block of triggers and actions. This principle is illustreted in the Fig. 4.

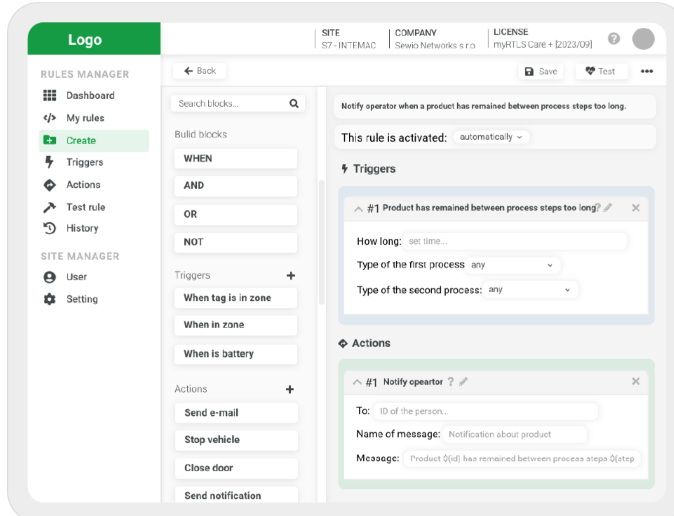


Fig. 4. User interface – create rule.

VI. Conclusion

In this paper we explored the possibilities of leveraging the principles of Low-code/No-code approach to industrial system logic. We aim to increase awareness of utilizing positional data provided by RTLS systems in inovative ways, making the most of the data potential towards the process of industrial digitization. The first part of the paper describes the rationale for creating rule-based engine systems over position data. The second part brings a summary of existing problems in the industry. The remaining sections describe the proposed system, its requirements, and the individual parts of the system.

Acknowledgment

Many thanks to the company Sewio Networks s.r.o. and their technical consultant Jakub Lešínský for his valuable feedback and assistance in the development of the rule-based engine.

References

[1] F. Zafari, A. Gkelias and K. K. Leung, ”A Survey of Indoor Localization Systems and Technologies,” in IEEE Communications Surveys and Tutorials, vol. 21, no. 3, pp. 2568-2599, 2019.

[2] S. Thiede, B. Sullivan, R. Damgrave, and E. Lutters, “Real-time locating systems (RTLS) in future factories: technology review, morphology and application potentials,” Procedia CIRP, vol. 104. Elsevier BV, pp. 671–676, 2021.

[3] R. Martins, F. Caldeira, F. Sa, M. Abbasi, and P. Martins, “An overview on how to develop a low-code application using OutSystems,” 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE). IEEE, Oct. 09, 2020.

[4] R. Sanchis, Ó. García-Perales, F. Fraile, and R. Poler, “Low-Code as Enabler of Digital Transformation in Manufacturing Industry,” Applied Sciences, vol. 10, no. 1. MDPI AG, p. 12, Dec. 18, 2019.

[5] KASS, Sebastian, Susanne STRAHRINGER a Markus WESTNER. Drivers and Inhibitors of Low Code Development Platform Adoption. In: *2022 IEEE 24th Conference on Business Informatics (CBI)* [online]. IEEE, 2022, 2022, s. 196-205 [cit. 2023-04-02]. ISBN 978-1-6654-6016-3. Dostupné z: doi:10.1109/CBI54897.2022.00028

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.“

Creation of a Digital Twin and Its Control Based on the ANSI/ISA-S88 Standard

1st Jan Holba
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech - Republic
xholba05@vutbr.cz

2nd Michal Husák
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech - Republic
xhusak08@vutrb.cz

***Abstract*—This paper deals with development of detailed digital twin of small-volume liquid storage unit, as well as its complete virtual commissioning. The original design of the existing unit is substantially modified to increase reliability and robustness. The control logic for PLC, according to ANSI/ISA-S88 standard, is also created. Based on this implementation of the control logic for the digital twin, its real counterpart will also be put into operation.**

***Index Terms*—Digital Twin, Industry 4.0, ANSI/ISA S88, Batch Process, Siemens NX**

I. Introduction

The aim of this work is to develop a detailed digital twin and fully commission it virtually. This digital twin is a counterpart of real small-volume liquid storage unit. The original construction of the existing unit will be substantially modified to increase its reliability and robustness. Additionally, control logic for the unit will be implemented based on the ANSI/ISA S88 standard for batch processes. Based on the implementation of the control logic for the digital twin, its real counterpart will be subsequently commissioned.

II. The Digital Twin

A digital twin (DT) is a complex model which allows detailed simulation of a real system. Thanks to the DT a system can be easily operated in virtual environment. This eliminates the problems associated with manipulation of physical system in the real world. DT cannot damage itself nor its surroundings. It does not wear down and does not depreciate itself in any other way.

During the life cycle of the machine, DT plays and irreplaceable role in development phase of the equipment. When creating the control logic for a given system, the programmer does not have to be located near the real machine. The given machine does not even have to physically exist yet. The control logic can be created remotely from the comfort of his home or office. The stress and fear of damaging expensive equipment is reduced, and the programmer can fully concentrate on his work. He can experiment as he pleases, restart the simulation at any time to any position, or pause it to debug his code. This is usually not possible with a real machine. After completing the control logic for the digital twin, close to none code modifications are necessary to get the real system up and running.

In addition to developing and testing of new concepts, the digital twin is also suitable for training new staff or for teaching. Workers can get familiar with the system in depth and learn how to react in the event of hazardous situation. During the training, the real machine performs its normal work and thus production is not limited in any way. In teaching, it is advantageous to use a DT mainly for financial reasons. Creation and virtual commissioning of a DT is significantly less expensive than development of fully physically functioning machine on which students can learn to code. The effectiveness of learning, whether in school or in industry, will be significantly increased, because the fear of the possibility of damaging a real, often expensive, machine is eliminated. Several people can use the simulation at the same time on different computers.

Simulation time can be scaled and action execution speed of the DT can be adjusted. It can be stopped at a certain point so necessary calculations can be performed. It can be run slowly for detailed analysis, or a large number of process cycles can be performed rapidly for measurement and analysis of long-term data without the need of lengthy measurements of a real system for several months. Likewise a huge number of different process modifications can be tested very quickly.

DT can be modified and optimized very easily based on the data analysis. During structure modification no material is used, so the equipment optimization is not expensive. The real system is then modified after thorough testing on the DT.

The DT simulation runs separately from real system. The complete feedback connection of the digital twin and the physical system results in the so-called Cyber-Physical System (CPS). The digital twin is constantly dynamically updated based on the measured values on the real machine. In this way, it is ensured that the real system and its digital twin move in a fixed manner and maintain a zero deviation even if the real system is delayed due to external forces or wear and tear, for example. However, the implementation of CPS is significantly more complex and will not be covered in this paper.

A. Digital Twin of Small-Volume Liquid Storage Unit

The digital twin that was created as part of this paper is a counterpart of real small-volume liquid storage unit. The task

DOI 10.13164/eeict.2023.57

of this unit is to fill a glass with a required mixture of liquids based on a prescribed recipe. This work builds on the work of Lukáš Rejchlík [1], but significantly modifies the original concept.

The central element of the cell is a three-axis manipulator. The vertical linear movement of the manipulator is driven by a BCI 6335 DC motor with a 21.6:1 gearbox. Horizontal linear movement is provided by a Hanpose 17HS4401 stepper motor with a NEMA17 flange. The same type of stepper motor also drives the rotary motion of the cup-grabbing arm. The actual position of the manipulator is being monitored by a pair of encoders and checked by limit sensors. The amount of remaining liquid in the storage unit is being monitored by the capacitive sensors.

An older version of the unit's 3D model already existed, but it was a classic, static 3D model, not a digital twin. For the creation of the DT, the individual parts of the model had to be converted from the original CAD software SolidWorks to the Siemens NX program and assembled. Subsequently, it was necessary to modify the structure of the cell and thus solve a number of problems.

The main problem originated from the method of dispensing liquids into the glass. A compression dispenser with an overflow chamber was used. The problem was that very high pressure needed to be exerted onto the mechanism to activate the dispensing. Since the handles of the dispenser as well as the manipulator itself are made of plastic using the 3D printing method, they could be damaged by the load. In addition, the overflow chamber of the dispenser has a constant volume, so it was not possible to dispense the exact required amount.

These problems are solved by the use of peristaltic pumps that allow us to dispense very precise volume of liquid. The lifespan of these pumps is typically quite long without the need of extended maintenance, they are also easy to control. However, pumps are no longer purely mechanical and need power to function.

The introduction of peristaltic pumps led to a modification of the original design of the unit. A substantial part of the unit was disassembled and redesigned. New bottle storage positions, that integrate pumps and capacitive sensors for liquid level monitoring, have been created. Because of the newly brought wires to each storage position, it was necessary to create a new, more compact counterweight housing that does not collide with the wires. Furthermore, the electrical wiring of the cell was modified. The final structure of the storage unit increases the number of storage positions from the original five to sixteen.

B. Physics Simulation and Virtual Commissioning of the Digital Twin

After completing the creation of the model, it is possible to switch from the Modeling application to the MCD (Mechatronics Concept Designer) application, which gives us the opportunity to create a very complex kinematic simulation. First, it is necessary to mark required parts as rigid and collision bodies.

Fig. 1. Digital Twin of the Storage Unit

Rigid bodies are objects that are subject to the laws of physics. They will be affected by external forces, such as gravity, or forces caused by other rigid bodies and actuators. Colliding bodies affect each other only in physical contact. These collisions can be unwanted impacts, e.g. collision of the manipulator with the structure of the cell. However, some collisions are necessary for the function, e.g. grasping the glass by the manipulator.

Rigid bodies can be connected to each other using various types of kinematic joints. Among basic joints are fixed, sliding, rotary etc. By assigning joints, we define the possible trajectory of objects. For example, a linear joint between a bearing and a guide rod ensures that the bearing will move along a precisely linear trajectory given the axis of the rod. If we want individual rigid bodies to be affected by the forces from other rigid bodies, we must assign them so called couplers. Example could be gear coupler for simulating gearboxes, or rack and pinion coupler for two-way transformation of rotary and linear motion.

Furthermore, it is necessary to simulate sensors and actuators. Sensors can work based on the principle of collision between two bodies. The calculation of collisions, however, tends to be computationally demanding, so it is appropriate to simulate the sensors based on the limit switch principle. Limit switch is activated if the tracked object exceeds the specified position, which we set to the sensor position. Speed control function has to be assigned to the drive units, and the maximum acceleration and torque has to be specified based on engine manufacturer's datasheet.

This digital twin is controlled by a PLC industrial control system. Communication is carried out using the OPC-UA server. The OPC-UA server reads the selected signals from the PLC and rewrites the values in the model. If minimal delay between PLC and simulation is required, it is advisable to use a PLC simulated by the PLCSIM Adv. program instead of a physical PLC and communicate directly with the model. In addition, the use of a simulated PLC will enable the complete development of control logic solely on a computer, without the need for any additional hardware. Beware, the S7-1200, unlike S7-1500, cannot be simulated using PLCSIM Adv. at the moment.

III. Testbed Self-Acting Barman

The created storage unit is part of the Self-Acting Barman project. The goal of the project is to create an automated production plant for mixed drinks based on the principles of Industry 4.0. The detailed description of the testbed can be found in the article An Industry 4.0 Testbed (Self-Acting Barman): Principles and Design [3].

IV. Standard ANSI/ISA S88

Industrial processes are divided into discrete, continuous and batch. The division is based on the nature of the final product. Batch processes are described by the ANSI/ISA-S88 standard. Output of the batch process are finite quantities of material (batches) [3]. Application of the standard is suitable for the Self-Acting Barman as the individual glasses represent separate batches of a product.

The standard is based on the division of batch processes into so-called models and their interconnection.

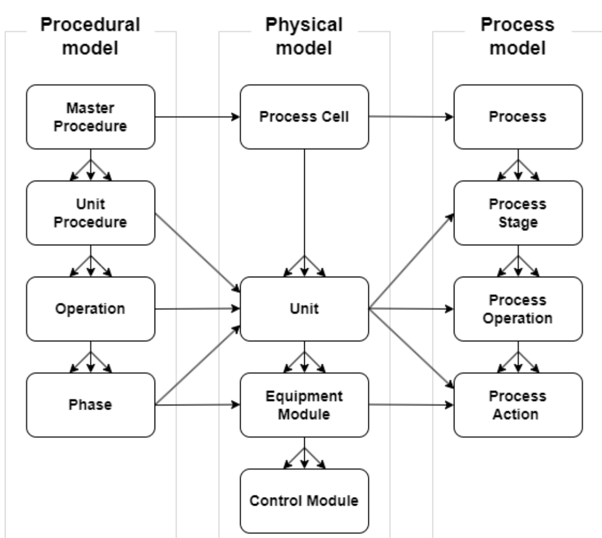


Fig. 2. ANSI/ISA S88 models [4]

The division into these models will allow a significant improvement in the synergy between technologists (process model), integrators (procedural control model) and assembly workers (physical model). Development of individual models, as well as layers, can take place in parallel and there is no need for an expert who understands all parts of the production process at the same time. The description of production technologies and procedures can be very complex. If we divide the process into groups of simple sections, it becomes significantly simpler and the whole process becomes modular. This makes it easily adjustable, and we can use individual modules repeatedly in the process.

The testbed Barman is defined as a process cell. As a complex production facility, it would not be possible to implement it without further division into process units. This paper deals with the partial goal of implementing a small-volume fluid storage unit.

A. 1.4. Process Model

A process model does not define a specific physical device. It describes the batch process from a recipe perspective. It defines It defines what raw materials, in what order and in what proportion need to be mixed, but does not specify what equipment needs to be used. We can compare the process model to a kitchen recipe. The model is structured into four levels and each level contains one or more parts.

In this case, on the highest level is the process that describes the entire workflow of drink preparation by the Barman testbed, from receiving the order from the user, through preparing the glass, pumping the drink and mixing it, to issuing the finished drink to the user.

In the lower layer, there are process stages that describe the work approach of individual process units. One of these units is the storage unit, which is the subject of this paper. The process stage of the unit is further divided into process operations and, if necessary, further into process actions. Process actions describe the approach to picking up the glass, moving it to the required positions, filling the required liquids and dispensing the glass.

B. Physical Model

The physical model describes the equipment that is needed to perform the procedure that is described in the process model. It describes specific physical equipment and divides it into individual sections. Combining physical equipment at lower level creates a new element at a higher level. This element then behaves as a whole, and a change at a lower level is possible only by redesign. The physical model is divided into seven layers. But the first three layers, i.e. Enterprise, Location and Operation, are not directly related to batch production and therefore will not be considered further.

The process cell is the entire testbed Barman. It consists of several process units. These are, as described in the article [3], a glass storage, a carbonated water maker, small and high volume liquid storages, a shaker and an ice crusher. Each process unit is further divided into equipment modules. For this storage unit, these modules are the transport module and the dispensing module. The equipment modules then consist of control modules such as stepper motor, DC motor, peristaltic pump, capacitive level sensor, etc.

C. Procedural Control Model

It combines the previous two models. It describes functions, that ensure the equipment, defined in physical model, is used in a way that leads to completion of the process defined in the process model. The model consists of four levels: master recipe procedure, procedures of individual units, operations and phases.

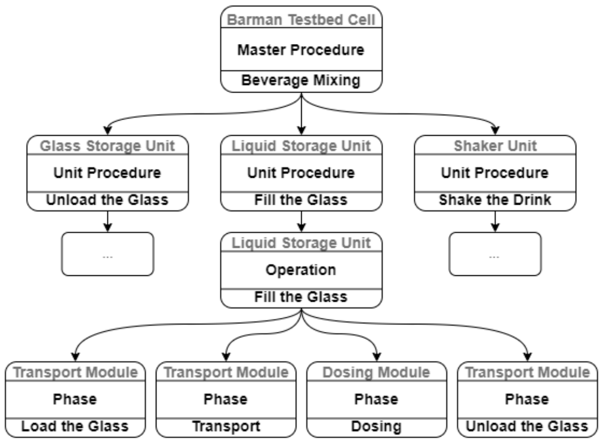


Fig. 3. Procedural Model for Barman Testbed

The master recipe procedure controls the entire process cell to perform the desired process. In the barman testbed, this master recipe procedure is to mix a drink and hand it to the user based on his request.

On a lower level are unit procedures that are related to individual units from the process cell. One of the process units of our testbed is my storage unit. The procedure of this unit describes all the functions required to dispense the desired liquids into the glass and pass it to the next unit.

The control logic is further divided into operations. The standard defines that only one operation can run on each unit at a time. It also says that a unit procedure consists of one or more operations. If there is no need to divide the production into several operations, there can be only one. Operations then consist of phases that form the lowest level of control logic. Several phases can run in parallel and can be repeated within one operation, as follows from the modular concept of the entire standard. Phases are the phases of loading the glass, transport, dispensing and unloading the glass.

The standard helps with structured control, where the entire process cell does not have to be controlled by one powerful and expensive control system, but the individual elements in the lower layers of the procedural model are controlled separately by simpler control systems. If we look at the Barman testbed, there is no single control system that independently controls the entire production, but each unit has its own control system with its own logic. Units communicate with each other using an administrative shell. The communication interface recommended by the S88 standard allows the use of a standard central MES for temporary development purposes until fully decentralized solution is available.

V. Testing

At the time of writing the article, the DT is fully operational virtually and work is underway to commission a real unit. Due to the fact that the DT is very detailed, the real commissioning is based purely on the calculations of conversions of the input and output signals of the control unit. At this point, the most important phase of transport is fully operational, which on a real cell performs the same movements as the DT. However these movements are not completely identical, because created simulation is classified as a DT rather than a CPS, as mentioned at the end of chapter II.

VI. Conclusion

The digital twin represents an ideal way of developing and testing new devices. It helps in development and production as well as in teaching. The digital twin of the small-volume liquid storage unit, which was developed as part of this paper, features complete kinematics simulation and is fully controllable by both a real PLC via the OPC-UA server as well as by a simulated PLC. The application of this digital twin enabled the remote development of the control logic from the comfort of home without the need to be present at the real unit, which was not yet complete at this stage of development. Furthermore, the digital twin significantly eased experimentation and debugging by eliminating fear of damaging expensive equipment, enabling full focus on control logic development thus increasing its efficiency.

A substantial part of the original structure of the unit had to be disassembled and redesigned. That led to increase in its reliability, robustness and accuracy of liquid pumping. At the same time, it was necessary to modify the electrical wiring of the unit and a complete electrical diagram was created.

The use of the ANSI/ISA S88 standard and the associated division into individual models and levels enabled the organized development of the control logic of the batch process. The created and debugged control logic require minimal modifications for real deployment.

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] L. Rejchlík, ”Návrh, konstrukce a programové vybavení inteligentního skladu pro testbed Průmyslu 4.0”, Brno : Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky, 2019. Diplomová práce. Vedoucí práce: Ing Václav Kaczmarczyk PhD.

[2] V. Kaczmarczyk, O. Baštán, Z. Bradáč, J. Arm, ”An Industry 4.0 Testbed (Self-Acting Barman)”: Principles and Design. 2018.

[3] ISA-The Instrumentation Systems, and Automation Society, ”ANSI/ISA-88.01-1995, Batch Control, Part 1: Models and Terminology”. 1995. ISBN: 1-55617-562-0.

[4] M. Husák, ”Návrh a realizace laboratorní úlohy na téma ”Standard ISA-S88”. Brno : Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2018. Bakalářská práce. Vedoucí práce: Ing Václav Kaczmarczyk, Ph.D.

Manufacturing system for Industry 4.0 demonstrator

Tomáš Martiník
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
xmarti95@vutbr.cz

Ing. Václav Kaczmarczyk, Ph.D.
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
kaczmarczyk@vut.cz

***Abstract*— This paper deals with MES systems, their general description, and a search of available open-source MES system solutions on the Internet. One of the found systems is chosen for extension and integration to demonstrator of the Industry 4.0 system. There are discussed some possible ways of extending the system.**

***Keywords*— MES systems, CRM, open-source, manufacturing ERP modules, EspoCRM, automation, Industry 4.0**

I. Introduction

Manufacturing companies today face increasing pressure to improve efficiency and reduce costs to remain competitive. To achieve these goals, many companies are turning to digital technologies and the Internet of Things (IoT) to create "smart factories" as part of Industry 4.0.

Industry 4.0 is the fourth industrial revolution, characterized by the integration of smart technologies and real-time data to optimize production processes and increase efficiency. One key aspect of Industry 4.0 is the implementation of standardized communication protocols and data exchange formats, which enable the seamless integration of different systems and equipment.

ISA-95 is a standard developed by the International Society of Automation (ISA) that provides a framework for integrating enterprise and control systems in manufacturing. By implementing ISA-95, manufacturers can achieve greater transparency and efficiency in their production processes by integrating different systems and exchanging data in a standardized way. [1]

One important system that can benefit from ISA-95 integration is the Manufacturing Execution System (MES). MES systems control a company's processes, materials, and resources to optimize production, increase productivity, and improve quality. In this bachelor's thesis, we will explore MES systems, their installation, setup, and use in the production process. [2][3]

This paper presents the process of researching available open-source MES systems, selecting one, extending it, and integrating it into an Industry 4.0 demonstrator system. The goal is to replace the current MES system in the Industry 4.0 demonstrator with a more robust and easy-to-extend system that includes additional functionalities and modules. By doing so, the benefits of using ISA-95 and MES systems in an Industry 4.0 context will be demonstrated, as well as how they can be used to improve efficiency, reduce costs, and increase competitiveness in manufacturing companies.

The current solution of MES system includes material, equipment, products, and process modules. It runs on a web server, however, it has been found difficult to extend this system because it's lacking documentation and it's not a robust solution.

II. Research

A. *Systems criteria*

The main criterion of researched systems is being open-source. The reasons are advantages of extending an open-source project – already created the core of the system (no need to create the system from scratch) that can be used both personally or commercially, possibly community or documentation, that will help developers to extend the system.

The intention is to research MES systems, but easily editable and extendable CRM (Customer relationship management) or ERP (Enterprise Resource Planning) systems may also be possible to integrate into the Industry 4.0 demonstrator system (after some transformation into a manufacturing system).

B. *Researched systems*

Two MES systems, that match our research criteria, were found. The first of them is Qcadoo. This system showed significant promise in terms of its features and capabilities, with several impressive features highlighted on the system's website. However, because of the lack of a complete installation pack for Windows, it was impossible to run and test this system.

The IMES system, on the other hand, was found to be easy to install but had a very limited range of functionality. The system lacked robustness and was not suitable for use in a large-scale manufacturing facility.

Both systems had some documentation, but for Qcadoo, the documentation is not very reliable. IMES documentation was intended for system users, not for developers. [4][5]

Three Enterprise Resource Planning (ERP) systems, that match our research criteria, were found.

First of them is Odoo - a comprehensive open-source ERP system that offers a wide range of modules, including manufacturing, inventory, and supply chain management. It is compatible with Windows 10, and its installation process is straightforward, with comprehensive documentation available on the system's website. The community support for Odoo is robust, with a large user community and active developer community. There are many manufacturing modules available for Odoo, including manufacturing planning, scheduling, and execution modules. [6]

ERPNext is another open-source ERP system that has the potential to be used as a MES system. The system's documentation is comprehensive, with user guides, video tutorials, and a forum for community support. The community support for ERPNext is active, with regular updates and bug fixes. While the system has some manufacturing modules, such as material resource planning and production planning, it lacks advanced MES features. The system is not available for Windows OS. Installation on Virtual Machine was impossible as the installation package wasn't complete. [7]

Dolibarr is an open-source ERP system that includes modules for CRM, accounting, and project management. It is compatible with Windows 10 and has a straightforward installation process, with documentation available on the system's website. The community support for Dolibarr is not as robust as Odoo or ERPNext, but it still has an active user community. Dolibarr has some manufacturing modules, such as inventory management and project management, but it lacks advanced MES features. [8]

One customer relationship management (CRM) system, that matches our research criteria, was found – EspoCRM. It is an open-source CRM system that has some potential to be used as a MES system. The system can be installed on a Windows 10 machine and has developer documentation available on the system's website.

While EspoCRM does not have advanced manufacturing modules, the system can be extended with custom modules and plugins. The administration, developer documentation, and community support provided by EspoCRM make it possible to add additional functionalities and transform them into a MES system with customized manufacturing modules.

C. *Research conclusion*

During my research, I wasn't able to run Qcadoo and ERPNext. Systems IMES and Dolibarr lacked robustness, advanced MES features, or ease to create new modules for manufacturing. Out of the six systems I investigated, it appears that the systems appropriate for transformation into MES systems are Odoo and EspoCRM. The decision of which system to choose ultimately depends on the company's and developers' specific needs and preferences, such as the features and functionalities they require, ease of use, scalability, and cost (as some features of these systems may be charged).

III. System selection and setup

After making research on various systems, I decided to use EspoCRM due to its many advantages, including its ease of use and ability to create custom modules. The system's backend is built using PHP programming language and the Laravel framework. The front end is built using JavaScript, specifically the Backbone.js framework. Additionally, EspoCRM uses a MySQL database to store data. [9]

To get started, a localhost server was set up on the computer and a MySQL database was installed. EspoCRM was then installed. The setup process was found to be relatively straightforward, allowing for quick initiation with the platform – a few minutes after downloading the source code, a user can open (run) the MES system in a browser.

Upon setting up the system, it already possessed several admin features such as user creation, role assignment, rights allocation, notification settings, and more.

IV. System customization

A. *Orders*

This module is designed to manage and track orders from creation to production execution. Orders have their database table, which contains details about their items (ordered products) and their quantities.

One of the key features of the Orders module is the ability to set priorities for each order. Users can assign a priority level to each order, which helps to ensure that critical orders are given priority during the production process. This feature can be useful in situations where there are limited resources or capacity constraints, as it enables users to make more informed decisions about which orders to prioritize.

B. *ISA-95 Object models*

Using the ISA-95 standard, we can define equipment classes with their respective properties. With this information, we can create equipment definitions and assign them to their corresponding classes. This approach allows us to have a clear understanding of the equipment's capabilities and limitations and enables us to assign values to equipment properties.

In the context of ISA-95 models, materials are also defined using classes, definitions, and properties. These models allow the creation of material definitions and the assignment of classes and properties to those definitions.

Additionally, the models allow for the tracking of material lots, which are specific batches of material that are used in the manufacturing process. By tracking the lot of material, it becomes easier to identify any issues or defects that may occur during the manufacturing process.

With the help of ISA-95 models, we can easily create material definitions and assign them classes with properties. This makes it easier to manage materials and their lots throughout the manufacturing process.

Using ISA-95 models, we can define personnel classes, which can be groups of personnel with similar job roles or functions. Each class can have its own set of properties such as skills, certifications, and qualifications. With this information, we can create personnel definitions and assign them to specific classes with their respective properties. This enables users to effectively manage personnel resources and assign them to appropriate tasks within the manufacturing process.

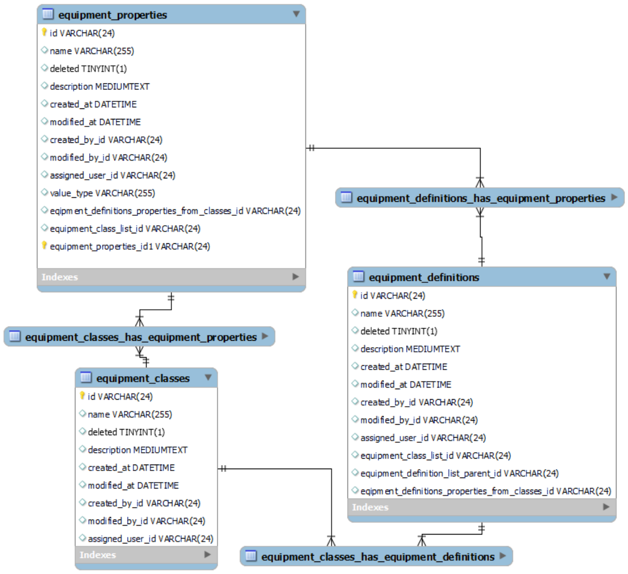


Fig. 1. Equipment object model (database)

C. *Product definition model*

Products are defined as a combination of materials and equipment used in a manufacturing process (product segment) to create the final output. Each product segment is defined by a product recipe, which lists the required materials and equipment.

The material segments define the type (definition or class) and quantity requirements of the materials needed to create the product. The equipment segments define the type (definition or class) and quantity requirements of the equipment needed to create the product.

D. *Process segment model*

In the ISA-95 models, process segments are defined as segments that connect material, equipment, and personnel to carry out specific manufacturing processes. These segments describe the steps involved in creating a product, such as mixing raw materials, assembling components, or packaging finished goods.

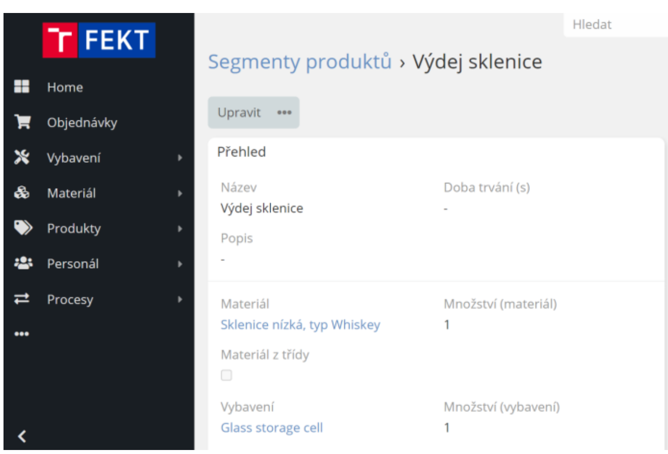


Fig. 2. Customized MES system

V. Conclusions and Future Work

A. *Conclusion*

Various modules and models were implemented into the new MES system. The Orders module allows efficient order management and prioritization, while the use of ISA-95 object models enables users to define equipment, materials, and personnel classes with their respective properties, making it easier to manage these resources throughout the manufacturing process.

Overall, the customization of the open-source system into a MES system provides a comprehensive solution for managing the entire manufacturing process, from order creation to product completion, with the ability to prioritize critical orders and efficiently manage resources. This system has the potential to significantly improve manufacturing efficiency and reduce costs, ultimately leading to increased profitability.

B. *Warehouse module*

The next step in this project is to create a warehouse module. This module will allow users to manage the inventory of materials and products that they have on hand. They will be able to track the quantity of each item, its location in the warehouse, and its movement in and out of the warehouse. This will help them to better manage stock levels and ensure that they always have the materials and products we need to keep our production process running smoothly.

C. *OEE module*

One of the next steps in the implementation plan is to develop an OEE (Overall Equipment Effectiveness) module for our MES system. The OEE module will allow us to track and measure the effectiveness of our production processes by analyzing data such as downtime, speed loss, and quality defects.

To implement the OEE module, we plan to leverage open-source chart libraries to display the OEE metrics in an easily digestible format. By doing so, we hope to provide real-time visibility into the performance of our manufacturing processes, identify constraints and optimize our production to maximize efficiency.

In addition to tracking the OEE, we plan to use the module to measure the effectiveness of our production processes. This includes analyzing data such as lead times, cycle times, and defect rates to identify areas where improvements can be made. By implementing an OEE module, we will be able to better understand our manufacturing processes and make data-driven decisions to improve overall efficiency and productivity.

D. *Use of AI*

The use of AI is being considered for potential implementation in this MES system. By using an open-source machine learning library like RubixML, we may be able to predict future demand for certain items based on historical data and adjust our inventory levels accordingly. This could help us to reduce waste and save costs in the long run. However, this is a more advanced feature that may require additional resources and expertise to implement, so it is a longer-term goal for the project. [10]

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] International Society of Automation: ISA95, Enterprise-Control System Integration [online]. Available at: https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95

[2] What is Industry 4.0—the Industrial Internet of Things (IIoT)?. Epicor [online]. Available at: https://www.epicor.com/en/blog/what-is-industry-4-0/

[3] MES center: MES system (Manufacturing Execution System) [online]. Available at: http://www.mescenter.org/cz/clanky/5-co-je-to-mes-system

[4] Qcadoo MES. GitHub [online]. Available at: https://github.com/qcadoo/mes

[5] IMES. GitHub [online]. Available at: https://github.com/jukbot/smart-industry

[6] Odoo [online]. Available at: https://www.odoo.com/

[7] ERPNext. GitHub [online]. Available at: https://github.com/frappe/erpnext

[8] What Dolibarr Does. Dolibarr ERP CRM Wiki [online]. Available at: https://wiki.dolibarr.org/index.php?title=What_Dolibarr_Does

[9] EspoCRM Documentation. EspoCRM Documentation [online]. Available at: https://docs.espocrm.com/

[10] RubiXML: Machine Learning for PHP [online]. Available at: https://rubixml.com/

Creation of Knowledge Base

1st Jiří Šiška
Department of Control and Instrumentation
FEEC, Brno University of Technology
Brno, Czech Republic
E-mail: xsiska17@vutbr.cz

***Abstract*—This paper is focused on designing a knowledge base in a field of ophthalmology for the NPS diagnostic expert system. Part of the paper is also to introduce the reader to the properties of knowledge systems, their characteristic features and their applications.**

***Index Terms*—knowledge base, expert system, artificial intelligence, knowledge engineering, ophthalmology, corneal ectatic disorders, keratoconus, keratoglobus, pellucid marginal degeneration**

I. Introduction

Expert systems are part of narrow artificial intelligence. By properly encoding the knowledge and experience of an expert (or experts) into a knowledge base, an expert system is able to make decisions at a highly expert level in a particular field. These systems are used, among others, mainly in medicine, agriculture, business, in the diagnosis of equipment malfunctions or in the field of psychology. They can be helpful not only in practice, but also in the education of students - in this case, for example, future ophthalmologists.

II. Expert Systems

Expert systems differ from other branches of artificial intelligence mainly by placing more emphasis on the quality of knowledge than on algorithms. The advantage of an expert system is transparency in decision-making. In contrast to artificial neural networks, where we do not know on the basis of which parameters the result was evaluated, with expert systems we can see exactly how the given answers influence the outcome of the consultation.

A. Properties

Distinctive features of expert systems include:

- The expert's knowledge is expressed absolutely explicitly, in the form of a knowledge base, which should be designed in such a way that the expert himself is familiar with it and is able to modulate it himself. The knowledge base must therefore have the possibility of a large degree of modularity.
- Dialog mode of obtaining data from the user. The expert system applies its knowledge to a particular case or problem provided by the user's responses.
- Ability to work with a degree of uncertainty. Similar to experts, an expert system must be able to work with uncertain answers (e.g. a description of the patient's subjective problems) and with uncertain knowledge.
- The drawn conclusion of the expert system is not dependent on a single parameter. The system must be able to provide advice even if the input data is incomplete. Thus, there must be multiple alternative ways to evaluate the hypothesis.
- The expert system is able to provide justification and explanation of the conclusions reached. If we want to replace the expert with an expert system, it is necessary for such a system to be able to explain the derivation process to the user in the same way as an expert.

The enumeration of these properties serves only to clarify the key ideas of expert systems and cannot be considered definitive. A particular system does not have to meet all of the points above in order to qualify as an expert system. [1]

B. Diagnostic Expert Systems

During the consultation, diagnostic expert systems are tasked with evaluating which of the previously given hypotheses is involved, based on the input data. They are the most common type of expert systems. They determine which hypothesis from a predefined number of target hypotheses best corresponds to the data related to a particular case. Thus, diagnostic expert systems work with a finite number of goals (hypotheses, diagnoses), from which they select the most suitable ones. [2]

The figure below shows a block diagram of the diagnostic expert system.

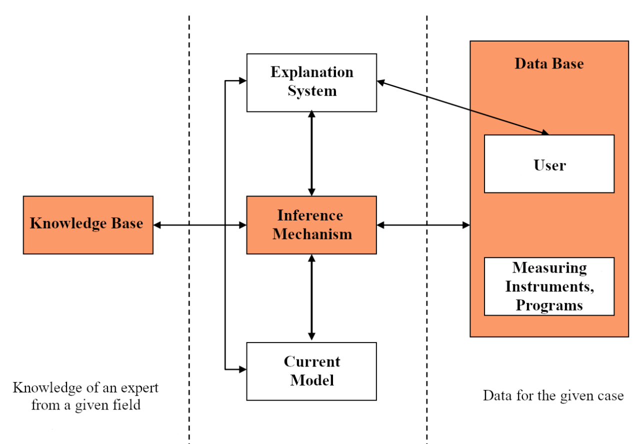


Fig. 1. A block diagram of a diagnostic expert system, *modified*. [2]

DOI 10.13164/eeict.2023.65

We can see that the system consists of three main parts which are highlighted orange.

- The knowledge base contains explicit, appropriately coded knowledge of the expert on the given issue, which is necessary for the proper functioning of the expert system.
- The data base represents data that we can obtain either from the user in the form of answers to asked questions or data read directly from measuring devices or programs. The data base affects the current diagnosis model.
- The inference mechanism is the core of the expert system. Based on the individual answers received from the data base, it refines the current model using the knowledge base.

When creating expert systems, efforts are made to make the program as modular as possible for future expansion and modifications. [3]

C. NPS Expert System

It is an expert system that was used for this application. The system is based on the original NPS32 computer application. NPS is a diagnostic expert system whose principle is based on rules. It is therefore a rule-based expert system. Today's form of the NPS web interface was developed by Ing. Lukáš Kořínek. It is based on the original NPS32 computer application and program implementation of the NPSCore computing core, which was originally developed without an user interface.

A slight disadvantage is the absence of an explanation system and the fact that user responses are the only possible source of the data base. However, the most fundamental advantage is the universality of the system, which is caused by a clearly defined syntax for knowledge base. [4]

III. Knowledge Base

The knowledge base is separate from the expert system. Until we provide a knowledge base to the expert system, it is an empty expert system. These systems are of much less value than if they contained a functional knowledge base.

First, it is necessary to be aware of the differences between the terms data, information and knowledge. Data is just filtered noise. We obtain information from data by selection, certain processing and assigning meaning. Knowledge is information that is thoroughly analyzed and organized in such a way that it can be used to solve a particular problem or to make a decision. We acquire knowledge through learning, experience, and interactions with the environment. There is also metaknowledge, which is knowledge about knowledge. [2], [4]

A. Knowledge Engineering

Expert systems have one major drawback, and that is the fact that an expert is needed to make such a system work properly. It is generally known that these gifted people, who devote a large portion of their lives to improving and gaining experience and knowledge in a certain field, are in high demand and do not exactly have an excess of free time. For this reason, it is often very difficult to design a functional and complete knowledge base. The field of knowledge engineering deals with the issue of obtaining knowledge from experts.

Knowledge engineering deals with filling expert systems with the most essential part, i.e. knowledge. In this field, there is therefore the greatest interest in techniques and methods of acquiring, formalizing, coding and testing knowledge. [1]

Great emphasis must be placed on the creation of the knowledge base, because the quality of the entire expert system depends on the quality of the knowledge. Therefore, both a knowledge engineer and an expert participate in this creation. The activity of a knowledge engineer can be summarized in several points listed below. [1]

- Identification of the problem - The knowledge engineer must first become acquainted with the issue and precisely formulate the problem.
- Design of the concept - After a deeper acquaintance of the knowledge engineer with the issue, the basic concepts and character of the data, with the active assistance of an expert from the area, the knowledge engineer can propose the conceptual model of the organization of relevant knowledge.
- Knowledge formalization - In this step, the knowledge engineer analyzes the conceptual model in terms of methods, techniques and tools. They choose to represent knowledge and formalize them appropriately.
- Implementation - The result of this phase of the knowledge engineer is a functioning prototype of the knowledge base.
- Testing and tuning - this step is the most time consuming, as it is constantly repeated in the cycle of testing, consulting results with experts and subsequent adjustment of the knowledge base.

The design of the knowledge base is not a linear process, on the contrary, we repeatedly return to the previous steps to iteratively approach the best solution. Each of these steps requires repeated meetings of a knowledge engineer with an expert, thus increasing the time demands. [1]

IV. Ophthalmology

The knowledge base in this paper is focused on the diagnosis of ectatic cornea diseases, which falls within the field of ophthalmology. Ophthalmology is a medical field dealing with the diagnosis and treatment of eye disorders and diseases. [6]

A. Corneal Ectasia

Corneal ectasia is produced by its thinning. More precisely, the stroma of cornea, which has the task of supplying the cornea sufficiently with water, is thinning. The stroma forms one of the corneal layers, consisting of fine collagen fibers and bounded by Bowman's membrane and Descemet's membrane. Patients suffering from ectatic corneal diseases are not recommended to undergo laser refractive procedures, due to the poorly predictable effect and the risk that the procedure will cause progressive deterioration of ectasia. [7], [5], [6] There are various types of corneal ectasia, such as keratoconus, keratoglobus and pellucid marginal degeneration.

B. Keratoconus

It is a clinical name of the state of the cornea, characterized by its gradual thinning and, as a result, its arching so that it gradually acquires conical shape. Progressive thinning and arching takes place most often in the paracentral part, less often in the central part of the cornea. [5]

It is one of the degenerative non-inflammatory cornea diseases with slow progression, which is manifested by increased corneal irregularity and deterioration of its optical properties, which often leads to a serious reduction in patient's visual acuity. It almost always occurs in each eye in different development stages. Keratoconus is classified by Amsler or Krumeich to four stages according to severity. The disease is most common in the second and third decade of life, when the greatest progression can also be observed. As aging progresses, the disease is generally stabilized, and after the age of forty years old, only deteriorates rarely. [5]

The low-degree keratoconus can be corrected by glasses or using soft or hard contact lenses. At the moment when there is further progression of arching, the patient should consider the possibility of available surgical procedures such as Corneal Cross Linking or corneal transplantation. [5]

C. Keratoglobus

It is a very rare, non-inflammatory, bilateral ectatic disease, which often occurs immediately after birth. Keratoglobus is usually not progressive or progresses only minimally. The disease is characterized by thinning and diffuse corneal protrusion, which leads to its typical arching of spherical shape. The treatment is developed for optimal correction using glasses or hard contact lenses. The uniform surgical procedure is not given due to the high risk of postoperative complications. [5]

D. Pellucid Marginal Degeneration

Pellucid marginal degeneration is a bilateral peripheral non-inflammatory ectasia of the cornea, which is characterized by the arch of its peripheral part, most often in lower quadrants. Compared to keratoconus, pellucid marginal degeneration begins at a later age, sometimes between the 20th and 50th years of life and progresses at a slower pace. It is manifested by irregular and non-correctional astigmatism. Treatment of this disease is complex, various types of special lenses or peripheral lamellar keratoplasty, or wedge keratectomy are considered. [5]

V. Knowledge Base Design

The knowledge base contains six main hypotheses and one additional hypothesis:

- Main Hypotheses
  - Keratoconus (first stage)
  - Keratoconus (second stage)
  - Keratoconus (third stage)
  - Keratoconus (fourth stage)
  - Keratoglobus
  - Pellucid Marginal Degeneration
- Additional Hypothesis
  - Normal cornea with astigmatism

The knowledge base consists of fourteen questions, four of which are conditional:

- Does the patient experience deterioration of vision?
  - Yes, significantly in one eye
  - Yes, significantly in both eyes
  - Yes, in one eye
  - Yes, in both eyes
  - No
  - Unknown
- What is the period of occurrence of problems?
  - 0 - 10 years
  - 10 - 20 years
  - 20 - 30 years
  - 30 - 40 years
  - Over 50 years and more
  - Unknown
- What is the refractive state of the eye?
  - Hypermetropia
  - Myopia
  - Unknown
- Occurrence of astigmatism?
  - Occurrence of higher astigmatism (greater than 2 D)
  - Occurrence of lower astigmatism (less than 2 D)
  - Irregular astigmatism
  - Unknown
- What does a topographic map look like?
  - Claws (moon)
  - Oval in the central part
  - Oval in the lower paracentral part
  - Circle
  - Hourglass (figure of eight)
  - Arching in full width
  - Unknown
- What is the thickness of the central part of the cornea?
  - Less than 200 μm
  - $200 - 300$ μm
  - $300 - 450$ μm
  - $450 - 530$ μm
  - More than 530 μm
  - Unknown
- Thinning of the corneal periphery?
  - Yes
  - Unknown
  - No
- Does the top of the cornea match the thinnest part of the cornea?
  - Yes
  - Unknown

- No
- What is the radius of curvature of the anterior surface of the cornea?
  - Less than 5.8 mm
  - 5.8 - 6.5 mm
  - 6.5 - 7.2 mm
  - 7.2 - 7.5 mm
  - More than 7.5 mm
  - Unknown
- What is the optical cardinality of the cornea?
  - Approximately 43 D
  - 45 - 49 D
  - 49 - 53 D
  - 53 - 55 D
  - 55 - 60 D
  - More than 60 D
  - Unknown
- Is there a Fleischer ring on the cornea?
  - Yes
  - Unknown
  - No
- Are there Vogt striae on the cornea?
  - Yes
  - Unknown
  - No
- Is there hydrops on the cornea?
  - Yes
  - Unknown
  - No
- Is there progression of corneal bulging?
  - Yes
  - Unknown
  - No

After answering these questions, the user should find out what kind of ectatic corneal disease it most likely is - the knowledge base is therefore successfully able to diagnose ectatic corneal diseases. However, it is possible this is not the final version of the knowledge base.

Due to the certain expertise of the questions and the overall focus of knowledge, the knowledge base is primarily intended for ophthalmologists who want to verify their decisions or for students familiar with this issue for the purpose of education.

It is also important to note that the functionality of the knowledge base was proven during testing with real data in cooperation with Mgr. Hana Řeháková. During this testing, questions were answered according to the subjects' actual measured parameters of the cornea.

VI. Conclusion

This paper describes the properties of expert systems. It discusses the diagnostic expert system and its internal structure in more detail. It also deals with an important part of the expert system, which is the knowledge base. It also describes the process of extracting knowledge from experts called knowledge engineering.

Other parts of the paper are focused on the topic for which I chose to implement the knowledge base and also deal with the design of the knowledge base itself. It describes the types of corneal ectasia, their characteristic features and mentions possible methods of treatment. It also contains hypotheses and questions together with answers that are used in the designed knowledge base.

The result of this paper is a functional knowledge base called *Rohovka*, which serves to determine the diagnosis of ectatic corneal diseases. This may not be the final version of the knowledge base, because at the time of writing this paper, the knowledge base is still being worked on. There also may be possible modifications of some questions and relationships between nodes based on the reevaluation of the concept of the knowledge base. The knowledge base require further tuning and testing to make the final knowledge base as factually accurate and useful as possible.

References

[1] V. Mařík, O. Štěpánková, J. Lažanský et al., Artificial Intelligence (2) [Umělá inteligence (2)], Prague: Academia, 1997, p. 15-137. ISBN 80-200-0504-8.

[2] V. Jirsík, "Expert Systems - BPC-UIN Lectures," Brno: FEEC BUT, 2021 [cit. 08. 12. 2022].

[3] V. Jirsík, "Expert Systems," [Expertní systémy], Habilitation thesis, Brno: FEEC BUT, 2003.

[4] L. Kořínek, "User Interface for an Expert System," [Uživatelské rozhraní pro expertní systém] , Bachelor thesis, Brno: FEEC BUT, 2019.

[5] P. Studený et al., Keratoconus [Keratokonus], Prague: Maxdorf, Jessenius 2020, edition MEDICA. ISBN 978-80-7345-665-8.

[6] Ophthalmology [online]. San Francisco (CA): Wikimedia Foundation [cit. 12. 28. 2022]. URL: https://en.wikipedia.org/wiki/Ophthalmology

[7] Stroma of cornea [online] [Stroma rohovky], Jindřichův Hradec: Čočky-kontaktní.cz, [cit. 12. 28. 2022].
URL: https://www.cocky-kontaktni.cz/slovnik/stroma-rohovky.html

Analysis and Detection of PWS Malware

1st Jan Blažek
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
exblazekja@vutbr.cz

2nd Jakub Křoustek
Department of Threat Research
Avast
Brno, Czech Republic
jakub.kroustek@avast.com

3rd Petr Dzurenda
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0002-4366-3950

***Abstract*—Cyberdefense became important, especially during the last decade. The rapid growth of information technologies caused a significant increase in cyber attacks and threats on the Internet. Malware analysis forms a critical component of cyberdefense mechanisms. In this article, we study the issue of malicious code and its various types, with a specific focus on the type known as PassWord Stealers (PWS). To do so, we deployed several methods of analyzing binary executable code, such as static and dynamic analysis, and sandboxing. We analyze 11 recently discovered malware families. From that, we discovered 3 new strains of malware, namely SevenStealer, NeedleDropper, and AtlantidaStealer. Furthermore, we have created appropriate detection rules for all of these malware, which have improved the detection capabilities of Avast anti-virus (AV) software worldwide. At the end of this article, we present the resulting data illustrating the spread of analyzed malware in the user base of the Avast company.**

***Index Terms*—YARA, Malware, Password Stealer, Reverse Engineering, Info Stealer, Static Analysis, Dynamic Analysis, Cyber Defence.**

I. Introduction

Nowadays, information technology is an inseparable part of human life. These technologies are used by people of all ages, backgrounds, and nationalities for daily activities such as ordering food, social interactions, and store payments. We trust these devices to serve our needs without considering the possibility of malicious activities happening in the background if our device is infected with malware. Malicious software has existed in the information technology industry from its beginning. Currently, there are tens of malware types and thousands of malware families attempting to infect our systems to take advantage of our data. With the growing technology industry and intensified digitalization, the threat of cyber attacks continues to rise. This fact opens up a number of questions. How can we protect our systems from being infected? How do they infect our PCs? And how can we classify the malware family that is trying to get into our devices? The main outcome of this paper is to create detection rules, which can be used to answer these questions.

A. Contribution and Paper Structure

This paper focuses on analysis and detection of either known and unknown info-stealaing malware. We analyze eleven malware families, which results in exposing malware's behavior and functions. In addition, we create YARA rules to detect the analyzed malware families. The rules focus on static sequences, behavior, network communication, and known named objects, which are exposed during the analysis. The malware families are analyzed both statically and dynamically using modern tools such as IDA Pro, x64dbg, and Cuckoo sandbox. During the analysis, the anti-analysis techniques implemented in malware's samples had to be overcome. Our detection rules led to improved detection mechanisms for more than 435 million users all over the world. In total, the detection rules blocked over 2 million attacks that tried to steal victims' data, mainly from browsers, mail clients, and crypto wallets. The main targeted geographical locations were Argentina, North America, and the middle of Europe. The detection rules created during this research are mainly focused on their usage in Avast's products. However, some of these rules can also be used by other malware detection solutions, thanks to the open-source YARA code.

The paper is organized as follows. Section II outlines the used terminology, Section III described malware families, which were analyzed during the research. Section IV presents the experimental tests and statistical analysis. In the last section, we conclude this work.

II. PRELIMINARIES

In this section, we outline the used terminology, i.e., what is malware, which are basic techniques for malware analysis, detection and classification, and how malware protects itself.

A. Malware and Password Stealers

Malware, short for malicious software, is a computer program, which performs malicious activity on victim's machine. Depending on the behavior, the malware is divided into several types, such as ransomware, info stealer, or trojan. A malware family refers to a set of similar samples within a specific type that are derived from the same source. PassWord Stealers (PWS), often called as info stealers, is the type of malware which gather data from an infected computer and sends it to the attacker. The most common targets are credentials, credit cards, crypto wallets, and user's documents. These data are extracted from specific application they belong to. We refer to [1] for more information about PWS.

B. Malware Analysis

The purpose of malware analysis is usually to provide the information needed to respond to network intrusions. The typical goal of this analysis is to determine what happened, obtain a list of the capabilities of the suspicious binary, or understand how to detect it in a protected network. The malware analysis being done by the following methods as defined in [2]:

- *Static analysis* - Consists of examining the executable file without viewing the actual instructions. This analysis is quick and can provide information about functionality, and produce file signatures.
- *Dynamic analysis* - This technique inspects the malware while running and observing its behavior in order to remove the infection, produce effective signatures, or both. This must be done in secure environment; otherwise, the analyst risks damaging their own network.
- *Sandboxing* - A security mechanism for running untrusted programs in a safe environment. Sandboxes often simulate network services to ensure that the binary being analyzed will run normally. After the execution ends, the sandbox usually provides a report about the behavior of the analyzed sample.

C. Malware Detection and Classification Techniques

The tools used for malware classification during this research is called YARA [3]. It is pattern matching tool, which is able to work with both textual and binary patterns. Each rule consists of set of strings and a boolean expression which determine its logic. The behavioral classification rules made during this research uses the cuckoo module to parse the output of the Cuckoo sandbox, where the malware is analyzed.

D. Malware Self-defense Mechanisms

Regardless of the application being developed, if end users are external to the developing organization and the software is not open source, it is advisable to incorporate anti-reversing measures into the program [4]. The malware usually implements several anti-reverse engineering techniques to make the analysis as hard as possible. As defined in [2], the basic approaches of anti-reversing are the following:

- *Anti-static analysis techniques* - They aim to minimize the amount of information that can be obtained from static analysis. A common way of achieving this is by applying obfuscation, which generally modifies the program's layout, logic, data, and organization in a way that maintains its functionality while making it less readable.
- *Anti-debugging techniques* - The purpose of these techniques is to discover if the malware is being debugged or not. There are several methods for accomplishing this purpose. The most common methods are: using the Windows Application Programming Interface (API), or manual anti-debugging checks being done via direct Process Environment Block (PEB) access, and code scanning.
- *Anti-disassembly techniques* - This set of techniques causes disassembling tools to produce incorrect results. These techniques are created by the malware authors manually. This is usually achieved by inserting special jump instructions which will confuse the disassembler, which will decompile unused code.

By armouring the code with anti-reverse engineering techniques, the level of skill and time required to analyze the malware increases. In this research, anti-reverse engineering methods had to be overcome in order to produce obtained results.

III. PWS Malware Analysis and Protection Mechanisms Design

In this section, we analyze 11 recently discovered malware families and create appropriate detection rules for them. SevenStealer, NeedleDropper, and AtlantidaStealer are new strains of malware that we discovered during this research and are described in more detail. Other malware families were also analyzed to create detection rules. We refer to [5] for more detailed analysis of described malware strains.

A. SevenStealer

We discovered SevenStealer on October 25th, 2022 while observing a set of malware samples that had only a general classification. This malware is written in GO and contains readable paths to GO packages that are used inside the program, which often contain the text "C_/Users/777/Desktop/stealer_v6/stealer". The name "SevenStealer" is derived from the reference to Winows user 777, who compiled the source code.

The malware steals data from several sources from infected computer such as browsers, web extensions, and crypto wallets. After reading the data from the targeted software, SevenStealer creates a zip file in memory using archive/zip GO package. Afterwards, the whole archive is encoded by the base64 algorithm and transferred over the network as a part of Hypertext Transfer Protocol (HTTP). POST request back to the attacker. The request is form-urlencoded, the zip archive is sent under B64 key, and current username is sent as the "Userid" HTTP header. Code transferring the file to the attacker is inside a *pkg.SendLog* namespace. Even the network communication is captured and the HTTP request transferring the data is recorded, it is not possible to directly extract the content of the zip file since the archive itself has an invalid format. The only way to read the contents of the file is by using the same library used by the malware to create the archive.

SevenStealer can be detected by multiple unique static sequences, which occur inside the Portable Executable (PE), such as, namespaces, array of ids of browser extensions, or sequences used during the communication. Also, as the malware uses only one Command and Control (C2) server, the malware is also detectable by the Internet Protocol (IP) address used during the communication.

B. NeedleDropper

We discovered NeedleDropped during an analysis of the FormBook sample. It was found that the FormBook sample had been dropped on the disk by another currently unknown malware. NeedleDropper is a self-extracting archive which contains tens of files. Most of the files are just random junk data pretending to be something useful. Searching for important information used by the malware is similar to searching needle in a haystack, and this is where the name NeedleDropper came from.

NeedleDropper is a complex malware that uses several legitimate files during its execution, including a Visual Basic script, a patched AutoIt interpreter, an AutoIt script, a configuration file in ini format, and a payload. Once the archive is extracted, it executes the Visual Basic script, which runs the AutoIt interpreter. The interpreter then reads the obfuscated AutoIt script and carries out the malicious activities found within it. The first action the AutoIt script takes is to read the configuration file, which specifies the behavior of the malware. This file contains several pieces of information, such as key file names used by the malware, extraction path, crypt key, and information about persistence. After reading the file, the malware reads the file containing the final payload and decrypts its contents by calling the *CryptDecrypt* Windows API function and using the key obtained from the configuration file. Finally, the malware starts *RegSvc.exe* in a suspended state and injects the decoded payload into its process.

During the analysis, it was discovered that NeedleDropper uses an -as-a-service business model and is used by threat actors to hide their payload. This malware is difficult to detect due to the large number of files contained within its archive, and its legitimate files help it evade detection. Global cybersecurity companies such as Avast [7] and IBM [8] have published articles about NeedleDropper to raise awareness and help protect against this threat.

The static detection rule created to detect NeedleDropper targets static sequences observed inside the AutoIt script and configuration file. Another way to detect this malware is by creating a behavioral rule that combines its file system write access and manipulation with *RegSvc.exe*.

C. AtlantidaStealer

AtlantidaStealer is the latest strain of malware that we discovered. Although we initially thought it was a new malware strain, during the analysis, we found similarities with BurmillaStealer, which is a MaaS information-stealing malware distributed by Passion Team. As of March 2023, the Passion Team is banned on the Russian forum where they advertised the malware.

AtlantidaStealer steals a significant amount of data from PCs, such as hardware information, information about installed programs, currently running tasks, browser data, Steam accounts, information from FileZilla software, screenshots, as well as documents from specific folders, such as the user's desktop, documents, and downloads. Installed programs are enumerated via a specific registry key. All the collected information is stored inside a zip archive, which is stored inside the RAM. To achieve faster execution time, the malware distributes the work into several threads that run asynchronously. The malware is obfuscated by an unknown obfuscator and checks the system for suspicious processes, loaded DLLs, filesystem locations, and system information that could indicate any kind of analysis being done. If the malware detects an analysis attempt, it pretends to crash with a runtime error. AtlantidaStealer has only one command and control (C2) server, which is used as a log storage. To distinguish which logs belong to which user, the malware contains a user ID value that is sent together with the zip archive back to the C2 server.

During the analysis, we found artifacts and similarities that indicated that AtlantidaStealer is not a new malware strain, but rather a newer version of BurmillaStealer. The behavioral detection rule focuses on a large number of anti-analysis techniques, together with a registry key access. The static detection rule aims for the SeviremAZ.exe static sequence, alongside other sequences representing used functions or classes.

D. Stealerium

Stealerium is an open-source info-stealing malware available on GitHub [6]. It offers several stealing capabilities to its users and can be fully customized by one of its components. This malware can be detected by the LegalCopyright property in the PE header, which refers to the malware's repository. Additionally, the malware can be classified by the crypt key and salt used by both the Stub and Builder.

E. OriginLogger

OriginLogger is the latest version of AgentTesla, one of the most notorious Malware as a Service (MaaS) information-stealing threats. During our analysis, we discovered that the stolen data, which is transferred as an HTML file, is not properly escaped, potentially leading to cross-site scripting (XSS) attacks. The most effective way to detect OriginLogger is by combining executed commands, registry key enumeration, and file system read accesses into a single behavioral rule.

F. OutSteel

OutSteel is an info-stealing malware written in AutoIt. It was discovered by the Computer Emergency Response Team of Ukraine (CERT-UA) [9] during a cyber war in Ukraine. To extract the AutoIt script from the binary, the Exe2Aut tool can be used. The static detection rule created for this malware strain focuses on static sequences visible inside the extracted AutoIt script. The rule aims to detect sequences used for the final bat creation and enumeration execution.

G. LdPinch

LdPinch is an information-stealing malware that primarily focuses on stealing data from browsers and FTP clients. It is predominantly active in Russia and the USA. The malware typically increases its privilege by setting SeDebugPrivilege. It is not usually packed, and thus it can be classified based on static sequences associated with its activities.

H. Siggen

Siggen is a trojan malware with info-stealing capabilities. It performs several layers of decryption before executing its final payload. Therefore, it is more effective to classify the malware based on its behavior.

I. EternityStealer

EternityStealer is a malware written in C# stealing data from well-known application. It can be detected via two C2 URLs or unique configuration sequences.

J. PandaStealer

PandaStealer is a C++ malware that primarily focuses on crypto wallets. During its execution, it decrypts the most important strings. Therefore, the detection rule to classify PandaStealer aims for a specific combination of used libraries, classified by strings, and the hash of the import table.

K. GraphSteel

GraphSteel is a malware written in GO, which steals information from a victim's machine. It accepts the address of the C2 server as an argument and communicates via open-source packages developed by Google.

IV. Experimental Results

In this section, we analyze effectiveness of created detection rules. The rules were firstly deployed into testing phase where they were tested against incoming samples without taking any action. After a certain amount of testing, the rule improved the detection of more than 435 millions of Avast's AV clients. The measurement lasted for approximately 4 months, from the date the rule was created until March 2nd, 2023.

A. Characteristics of the Analyzed Malware

Based on our analyses, most malware samples targeted credentials from FTP, mail, and browser clients. On the other hand, the least observed functionality was instant message (IM) client enumeration, together with keyboard sniffing. Almost all malware samples implemented some kind of anti-analysis technique, typically obfuscation, Virtual Machine (VM) detection, and debugger detection.

B. Effectiveness of Created Detection Rules

Our detection and classification rules were deployed in Avast's production system. Firstly, each rule was deployed as a test and was promoted into production after a certain level of testing. Table I illustrates the number of unique hits per malware family.

TABLE I
THE NUMBER OF UNIQUE HITS PER MALWARE FAMILY. (EFFECTIVE DATE: 2ND OF MARCH 2023)

Strain	First Rule Deployment Date	Number of Hits
SevenStealer	29-10-2022	102
Stealerium	15-09-2022	1 941
NeedleDropper	26-10-2022	2 225
BurmillaStealer	16-01-2023	16
OriginLogger	24-10-2022	9 382
OutSteel	03-10-2022	6
LdPinch	04-09-2022	2 204
Siggen	03-12-2022	1 466
Eternity Stealer	26-12-2022	858
PandaStealer	20-10-2022	316
GraphSteel	16-10-2022	6

C. Blocked Attacks

After certain amount of testing, detection rules were promoted into the production to improve detection mechanism of Avast's AV clients all over the world. Only detection rules focusing on OutSteel or GraphSteel were not promoted due to low amount of incoming samples for this family. The amount of blocked attacks per malware strain is depicted in Table II.

TABLE II
THE NUMBER OF BLOCKED ATTACKS TARGETING END USERS. (EFFECTIVE DATE: 2ND OF MARCH 2023)

Strain	First Rule Deployment Date	Blocked Attacks
SevenStealer	14-11-2022	480
Stealerium	24-09-2022	4 903
NeedleDropper	29-10-2022	42 005
BurmillaStealer	01-02-2023	1 028
OriginLogger	28-10-2022	2 428 955
LdPinch	24-11-2022	3 612
Siggen	28-12-2022	2 308
EternityStealer	14-01-2023	3 209
PandaStealer	25-10-2022	682

V. Conclusion

In this work, we analyzed the functionality of 11 malware families and created several YARA detection rules capable of identifying them. The research presented in this paper has contributed to a better understanding of modern malware threats and their impact on global cyber security companies. The detection rules have improved the detection capabilities of over 435 million AV clients worldwide and have blocked over 2 million attacks targeting these users.

Acknowledgment

We would like to thank Avast company for providing access to its infrastructure and resources during this research.

References

[1] *Info stealers* [online]. Malware Bytes [cit. 2023-03-01]. Available: https://www.malwarebytes.com/blog/threats/info-stealers

[2] Sikorski, Michael, and Andrew Honig. *Practical Malware Analysis: A Hands-On Guide to Dissecting Malicious Software.* San Francisco: No Starch Press, 2012. Print.

[3] ALVAREZ, Victor. *Yara* [online]. GitHub [cit. 2023-03-01]. Available: https://virustotal.github.io/yara/

[4] Eilam, Eldad, and Elliot J. Chikofsky. *Reversing: Secrets of reverse engineering.* Indianapolis, IN: Wiley, 2005. Print.

[5] BLAŽEK, Jan. Analysis and Detection of PWS malware [online]. Brno, 2023 [cit. 2023-03-07]. Available: https://www.vutbr.cz/studenti/zav-prace/detail/151178. Bachelor's thesis. University of Technology in Brno, Faculty of Electrical Engineering and Communication Technologies, Institute of Telecommunications. Head of work Petr Dzurenda.

[6] *Stealerium* [online]. GitHub, 2022 [cit. 2022-11-27]. Available: https://github.com/Stealerium/Stealerium

[7] *NeedleDropper* [online]. Avast, 2023 [cit. 2023-02-05]. Available: https://decoded.avast.io/threatresearch/needledropper/

[8] *NeedleDropper - A New Strain Of Dropper Malware* [online]. IBM, 2023 [cit. 2023-02-05]. Available: https://exchange.xforce.ibmcloud.com/threats /guid:6316e7de70958528c5b2dc3c8539f4d5

[9] *Cyber attack of the UAC-0056 group on state organizations of Ukraine using SaintBot and OutSteel malware (CERT-UA 3799)* [online]. CERT-UA, 2022 [cit. 2022-12-19]. Available: https://cert.gov.ua/article/18419 (author's translation)

Photovoltaic Battery Charger with Sun Tracking

Martin Knob, Jan Kufa

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

E-mail: 221054@vut.cz, kufa@vut.cz

***Abstract*— This paper describes the design of a photovoltaic battery charger. The charger can track the sun and tilt the panel in azimuth and elevation. The device includes an Maximum Power Point Tracking (MPPT) regulator to maximise the energy from the photovoltaic panel. This photovoltaic charger is designed for Li-ion batteries and is equipped with appropriate battery protection. The system is battery-powered and therefore energy self-sufficient. Another purpose of the device is to measure accurate data for further processing. It will measure the energy consumption required for its operation and the energy produced by the photovoltaic panel. The device communicates wirelessly via Wi-Fi and sends the measured data to the Thingspeak server.**

Keywords—Photovoltaic panel, tracker, MPPT regulator, battery, charger, microcontroller, construction

I. INTRODUCTION

This thesis deals with the design of a photovoltaic battery charger with sun tracking. The photovoltaic (PV) panel is tilted in azimuth and elevation. This was done to optimise the performance of the monocrystalline photovoltaic panel. This type of panel has the highest theoretical efficiency, but only if direct sunlight falls on it.

The charger is powered by a 18650 Lithium-ion (Li-ion) battery, so it doesn't need to be connected to the electrical power grid and is energy self-sufficient. It charges the same type of battery using the integrated circuit LT3652HV [1]. The charger is equipped with appropriate charging circuit and battery protection. The battery is charged by the Constant Current followed by Constant Voltage (CC-CV) method. The CC-CV method first charges the battery at a constant current, and when a certain voltage level is reached, it starts charging at a constant voltage.

The device includes an MPPT regulator to maximise the energy from the PV panel. The regulation method used here is Constant Voltage. The device is controlled by an ESP32 microcontroller and sends the measured data via Wi-Fi [2]. The evaluation of the position of the sun is done using voltage dividers with photo resistors. The voltage levels on each divider are compared. After measuring the individual voltage levels, the position of the PV panel is changed in the right direction. The PV panel is moved by two Direct Current (DC) motors attached to the tracker structure.

Solar tracking photovoltaic panels are rare on the market. Most of the time, they are large devices or small kits. The advantage of this proposed system over commercially available devices is that we can measure the output at many locations, so we know exactly what is happening with the device. Another big advantage of this system over other devices is the wireless data transmission, so we can remotely monitor all measured quantities.

II. CHARGING CIRCUIT

The LT3652HV integrated circuit from Analog devices was selected for the charging circuit. It is a circuit designed for solar applications with an integrated MPPT regulator [1]. It uses a constant-voltage method that compares the panel voltage with a voltage reference. The diagram of this MPPT and charging circuit is shown in Fig. 1. The voltage reference is set by a voltage divider. The voltage divider has been implemented by a potentiometer, due to the possibility of changing the voltage reference value.

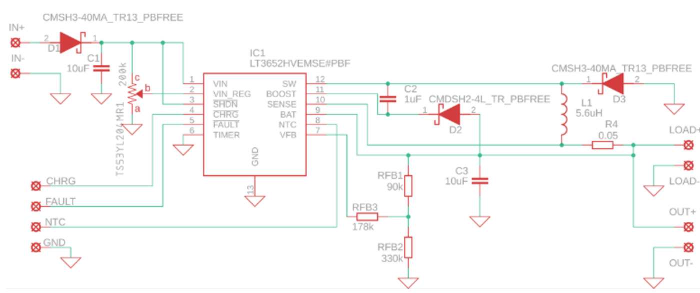


Fig. 1. Diagram of MPPT and charging circuit

The charger uses the CC-CV charging method. This method first charges at a constant current and when a defined battery voltage level is reached, it starts charging at a constant voltage. CC-CV is used to charge lithium batteries. This charger is charging a Li-ion battery, so this charging method is suitable. The maximum charge current is limited to 2 A. The charge voltage for this circuit is set by a voltage divider. In this case the charge voltage is set to 4.2 V.

Charging termination for this circuit can be done by time termination or C/10 termination. C/10 termination means that charging is terminated when the current value falls below 1/10 of the programmed maximum. After this termination, the charger enters a standby mode, where it only supplies 85 µA of current and restarts charging when the battery drops by 2.5%. The Printed Circuit Board (PCB) was designed in Eagle and is shown in Fig. 2.[1]

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.

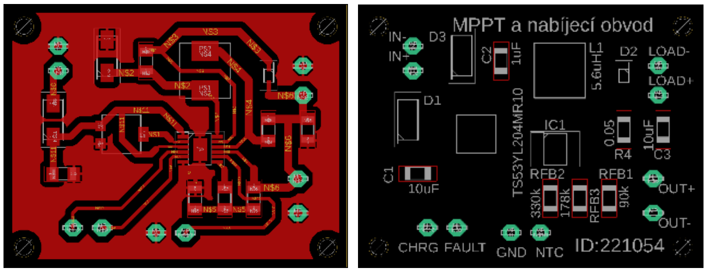


Fig. 2. PCB board of MPPT and charging circuit

III. Tracker Construction

The main requirement for the design was movement in azimuth and elevation. The structure was designed for a small 6 W monocrystalline photovoltaic panel with dimensions of 200 mm x 170 mm. A large gear ratio of 12:1 was chosen to achieve the required sensitivity and to avoid oscillation of the tracker. The tracker construction was designed in Fusion 360 and can be seen in Fig. 3. The design is produced on a 3D printer using Polyethylene Terephthalate Glycol (PETG) material [3]. The PETG material was chosen for its strength, good printability and temperature resistance. For outdoor use, temperature resistance is important as the structure will be exposed to direct sunlight.

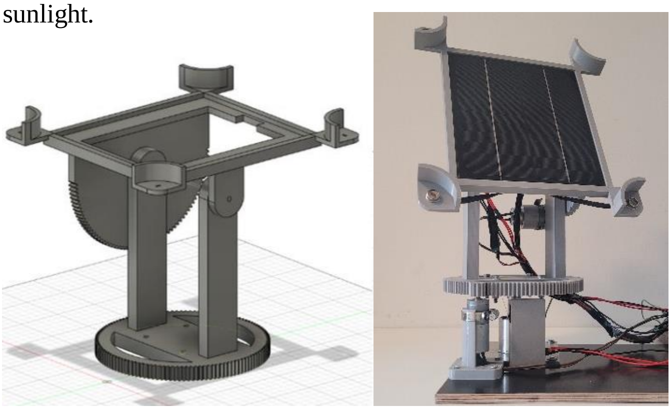


Fig. 3. Design of tracker (left), prototype of the Sun tracker (right)

The PV panel is at the top of the tracker. In the corners there are brackets for photoresistors that will sense the position of the sun. Shields are placed on the photoresistor holders to better evaluate the position of the sun. The motor holder is shown in Fig.4. The motor holder for the elevation movement is fixed to the structure. The bottom motor holder is bolted to the support board.

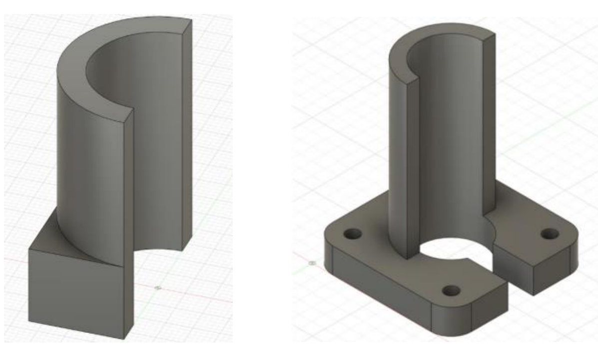


Fig. 4. Top motor holder (left) and bottom motor holder (right)

The printed tracker structure is bolted to the support construction. The support construction is shown in Fig.5. The support construction is made of stainless steel and includes two bearings for stability to prevent the structure from bowing. The support construction is bolted to the support board. All other components such as the microcontroller, MPPT, charging circuit and other modules are also mounted on the support board, but are hidden in a suitable box to protect them from external influences.

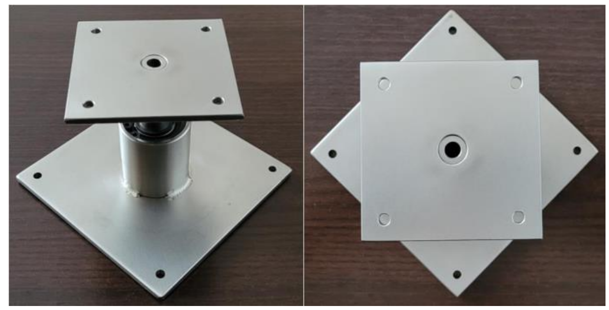


Fig. 5. Support construction

IV. Concept of Photovoltaic Battery Charger with Sun Tracking

The requirement for the charger was to track the sun in 2-axis. Two DC motors with gearboxes were chosen as the drive. Position sensing is performed by voltage dividers with photoresistors. The photoresistor changes its resistance depending on the number of incoming photons. The voltage levels are compared by the ESP32 microcontroller [2]. Based on this data, the microcontroller instructs the L298N module to move the motors [4]. The L298N module is an H-bridge that allows the DC motors to be controlled in both directions. There is a 5 V rectifier built into the L298N module, so the ESP32 microcontroller is powered directly from it.

The tracker is powered by a Li-ion battery with a nominal voltage of 3.6 V. Because we need to supply 8 V to the L298N module we need to use a DC-DC step-up converter. A voltage and current sensor INA3221 is connected to the battery to measure the power consumption [5]. This module is used to detect the power consumption of the tracker.

The next part of the system is a monocrystalline 12 V PV panel. The voltage and current sensor INA3221 is connected to the output of this panel [5]. The purpose of this module is to detect the amount of energy produced by the PV panel. The energy produced by the PV panel is connected to the MPPT and charging circuit. This charging circuit uses the CC-CV method to charge a Li-ion battery with a nominal voltage of 3.6 V. This charging circuit is described in detail in Chapter II. The output of the charging circuit is again measured by the INA3221 module to determine the losses in the charging circuit [5].

All components have been selected to be energy-efficient. To save energy, a real time clock has been added to this device. Based on this clock, individual power measurements and evaluation of the sun's position will be performed at intervals. When measurement is not in progress the microcontroller will be in sleep mode. The microcontroller will also be in sleep mode at night. The microcontroller will send the measured data to the Thingspeak server [6].

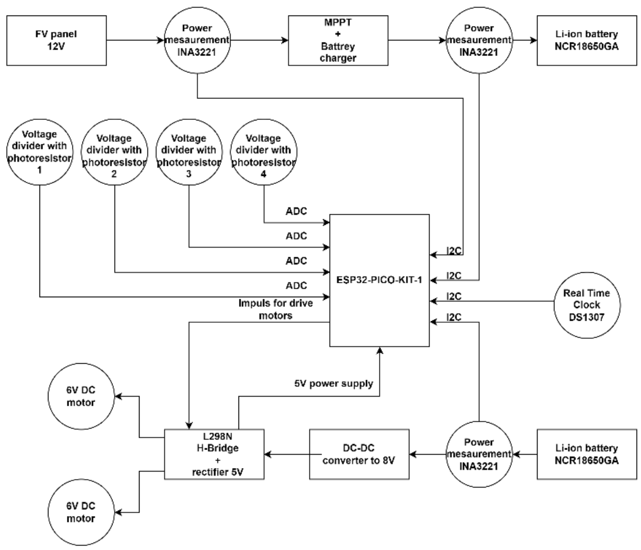


Fig. 6. Block diagram of PV battery charger with sun tracking

A. Photovoltaic Panel

There are many factors to consider when choosing a photovoltaic panel. We must consider the individual requirements of the application. We need a PV panel that is compact in size and low weight so as not to overload the structure. As the tracker will track the sun in both azimuth and elevation, a monocrystalline PV panel seems to be the best choice. This type of panel has the highest theoretical efficiency when exposed to direct sunlight. For this reason, a monocrystalline panel from Eclipsera was chosen, which weighs only 120 g and has dimensions of 170 mm x 200 mm [7]. This panel has a nominal output voltage of 12 V and a maximum current of 0.5 A.

B. Battery

This photovoltaic battery charger with sun tracking contains two identical batteries. One battery powers a circuit with a microcontroller to track the sun and measure the data. The second battery will be charged. The batteries used are Panasonic 18650 Li-ion batteries. These batteries have a nominal voltage of 3.6 V. They were chosen for their universality. A charging circuit has been designed for these batteries and is described in detail in Chapter II. [8]

C. Microcontroller

The microcontroller in this device evaluates the position of the sun, controls DC motors, measures electrical values and sends the measured data wirelessly to the Thingspeak server. For this reason, the ESP32 microcontroller was chosen because it includes Bluetooth and Wi-Fi [2]. Voltage dividers with photoresistors and other modules are connected to the microcontroller. The block diagram of this can be seen in Fig.6. In this block diagram we can see what is connected to the microcontroller.

D. Electric Motor and Control Modul

The tracker can be driven by DC motors, stepper motors or servo motors. Finally, 6 V DC motors 25GA-370 with gearbox and rotation speed 26 Revolutions Per Minute (RPM) were chosen [9]. The low speed is chosen here because of the movement is always only a small step. Combined with the 12:1 gear ratio on the structure, the desired speed should be achieved and the movement should not be too fast, so better sensitivity is achieved and the structure should not oscillate. Because the sun will only change its position a little bit in a whole day. DC motors were chosen for their simplicity and price. The control of these motors is implemented by the L298N module, which is basically an H-bridge with an integrated 5 V rectifier, so it can be used as a power supply for the microcontroller [4]. We supply a voltage of 8 V to the L298N module because there is a voltage loss of 2 V on the module itself. This ensures that the motors are supplied with 6 V and the microcontroller with 5 V.

E. Sun Tracking

The position sensing of this device is realised by voltage dividers using a photoresistor. The photoresistors used are Token PGM1205-MP with a resistance of 140-300 kΩ at 10 LX light [10]. The photoresistor with the highest resistance was deliberately chosen because most photoresistors are designed for measurements in lower light conditions. In direct sunlight the irradiance is higher. The divider is supplied with a 3.3 V DC voltage and the microcontroller compares the voltage levels of the individual photoresistors. If the voltage levels are equal, the PV panel is tilted at right angles to the sun. If the voltages on the photoresistors are different, then the panel must be tilted in the appropriate direction.

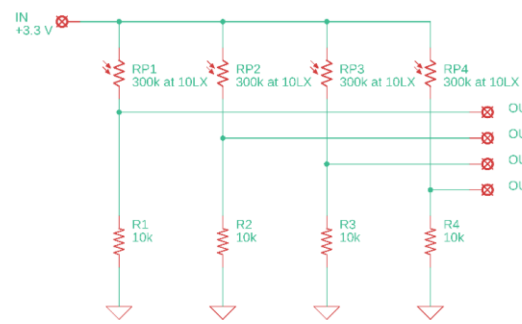


Fig. 7. Circuit diagram of voltage divider with photoresistors

The microcontroller always averages two neighbouring photoresistors to eliminate inaccuracies. The top, bottom, left and right are averaged together. Which photoresistors are averaged is shown in Fig.8. Based on this data the microcontroller will instruct the motor to move.

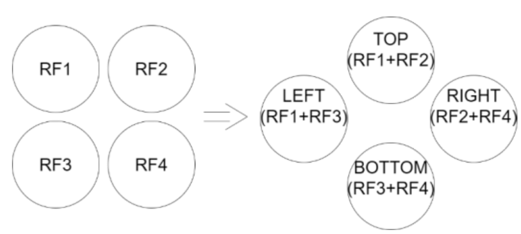


Fig. 8. Averaging of photoresistor

F. Other Modules Used

- Voltage and current sensor INA321:
  This is a three-channel voltage sensor that communicates with the microcontroller via I2C. This module measures voltages from 0 V to 26 V and the maximum current it

can measure is up to 1.1 A. The limitation of this current can be modified by changing the resistance of the shunt to obtain a higher sensitivity or, in the opposite case, to measure a larger current range. This module is used in the proposed device to measure the power supplied by the PV panel, the power to the battery to detect the losses on the charging circuit and the power required to operate the tracker. [5]

- Real time clock (RTC) DS1307:
  This is a module that provides real-time information to the microcontroller. It communicates over the I2C and has a backup power supply, which is good, for example, when the battery that powers the microcontroller and other devices runs out. The time information remains correct after the battery is replaced by the power supply system of the tracker and the device can continue to operate without further setup. The time information is used by the proposed device to wake up and put to sleep the microcontroller to save power. [11]

- DC/DC step-up converter SX1308:
  This is an adjustable step-up converter, which means that it changes the DC voltage from a smaller value to a DC output voltage of a larger value. The input voltage of this module is 2 V to 24 V and the output voltage is 2 V to 28 V. The maximum output current of this circuit is 2 A and the efficiency is 95%. This module is used to boost the 3.6 V voltage from the Li-Ion battery to 8 V for the L298N circuit. [12]

G. Communications

Communication with other devices is wireless. A mobile phone or a laptop can be used for wireless communication with the microcontroller. The ESP32 microcontroller is connected to a wireless network [2]. The microcontroller sends the measured data to the Thingspeak server, where we can access it. We can also connect to the device via Bluetooth, for example with a mobile phone, to access the measured data. The microcontroller communicates with the INA3221 measurement modules via I2C [5]. For the wireless connection, the ESP32 microcontroller was chosen because it includes Wi-Fi and Bluetooth.

V. Conclusion

The goal of this work was to design a system that would charge a battery using a PV panel. This system would actively track the sun and rotate the panel in azimuth and elevation to maximise the efficiency of energy conversion from the sun. The option of using either a monocrystalline or polycrystalline panel was offered, with the monocrystalline panel ultimately being chosen due to its higher efficiency, as the panel is still oriented perpendicular to the sun and shows more of the difference between a static design and an active tracking design. A Li-Ion 18650 battery was chosen for its universality, capacity and other characteristics, such as the absence of memory effect, which means that we can recharge it at any time, regardless of its previous charge or discharge level. There are two of these batteries used in the device, one as a power supply for the microcontroller and other components used, due to energy self-sufficiency so that no external power source needs to be connected and the other battery is charged from the PV panel. For these batteries, a charging circuit was designed that uses the CC-CV charging method, which means that it charges first with constant current and then with constant voltage. The LT3652HV integrated circuit was chosen for this charging circuit, which can also perform MPPT control using the constant voltage method. Low self-power has been emphasised in the design. An ESP32 microcontroller was chosen for the control, which will evaluate the position of the sun using photo resistors. It also instructs the DC motors to move in the correct direction to keep the panel perpendicular, and measures and sends the measured data to the Thingspeak server. Another purpose of the device is to measure accurate data for further processing. The purpose of the device is to determine if it is worthwhile to do sun tracking with a small 6W panel. The goal is to see if the device produces more energy than it needs to operate. Measurements are being taken at multiple locations to see if such a system makes sense for a small panel or if it is only worthwhile for larger PV panels.

Link to the video with Photovoltaic Battery Charger with Sun Tracking: https://youtu.be/6b5pFmNWNhs

References

[1] LINEAR TECHNOLOGY. LT3652HV. In: Analog Devices [online]. [cit. 2022- 12-20]. Available from: https://www.analog.com/media/en/technicaldocumentation/data-sheets/3652hvfb.pdf

[2] ESPRESIFF: ESP32-PICO-KIT-1 [online]. [cit. 2023-03-29]. Dostupné z: https://docs.espressif.com/projects/esp-idf/en/latest/esp32/hw-reference/esp32/get-started-pico-kit-1.html

[3] Autodesk: Fusion 360. Autodesk: Fusion 360 [online]. [cit. 2023-03-29]. Dostupné z: https://www.autodesk.cz/products/fusion-360/overview?term=1-YEAR&tab=subscription

[4] ECLIPSERA. Driver pro krokové a stejnosměrné pohony L298N [online]. [cit. 2023-03-29]. Dostupné z: https://dratek.cz/docs/produkty/1/1271/1420490399.pdf

[5] ECLIPSERA. Senzor napětí a proudu s I2C sběrnicí [online]. [cit. 2023-03-29]. Dostupné z: https://dratek.cz/docs/produkty/0/708/1523338818.pdf

[6] ThingSpeak: ThingSpeak for IoT Projects. ThingSpeak: ThingSpeak for IoT Projects [online]. [cit. 2023-03-29]. Dostupné z: https://thingspeak.com/

[7] In: ECLIPSERA. : Solární panel 12V 6W až 500mA s DC konektorem [online]. 2022 [cit. 2022-12-18]. Available from: https://dratek.cz/docs/produkty/1/1691/1643025601.pdf

[8] In: PANASONIC. : Specifications for NCR18650GA [online]. [cit. 2022-12-18]. Available from: https://www.tme.eu/Document/ae3467edeab6882567a6fea9726e74dc/NCR18650 GA.pdf

[9] Dratek.cz: Motor s převodovkou 25GA-370 6V DC 26RPM. Dratek.cz: Motor s převodovkou 25GA-370 6V DC 26RPM [online]. [cit. 2023-03-29]. Dostupné z: https://dratek.cz/arduino/121016-motor-6v-dc-26rpm-s-prevodovkou-silny-tocivy-moment-25mm.html

[10] In: TOKEN ELECTRONICS INDUSTRY. Token: Light-Dependent Photoresistors for Sensor Applications. [online]. [cit. 2022-12-19]. Available from: http://token.com.tw/pdf/resistor/cds-resistor-pgm.pdf

[11] Dratek: Hodiny reálného času DS1307 [online]. [cit. 2022-12-20]. Available from: https://navody.dratek.cz/navody-k-produktum/hodiny-realneho-casu-ds1307.html

[12] ECLIPSERA MODULES. Step-UP měnič 2A nastavitelný SX1308 [online]. [cit. 2023-03-08]. Available from: https://dratek.cz/docs/produkty/1/1079/1458853583.pdf

Modular Honeypot for IT and OT

Martin Nečas
Department of Telecommunications
The Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xnecas29@vutbr.cz

***Abstract*—This paper presents a design and development of a modular honeypot for information and operation technology (OT) that can be easily expandable and scalable. The honeypots send captured data to a web server. The honeypots can capture and archive all communication or specific data. The measured data are processed and shown to the user via a user-friendly website.**

***Index Terms*—containers, honeypots, IT, network monitoring, OT**

I. Introduction

This article focuses on designing modular honeypots for Information Technology (IT) and Operational Technology (OT) environments, which can enhance cybersecurity in critical infrastructure.

Several OT protocols lack the security to encrypt or authenticate communications. Consequently, they are exposed to various cyberattacks, such as virus infections and denial of service attempts [1]. In addition, many OT systems use outdated hardware that is challenging to secure and does not have the newest security features. Critical infrastructure systems which are susceptible to cyberattacks could lead to hinder or even endanger public safety. Some OT protocols started slowly implementing Transport Layer Security (TLS) for security propose. It is worth noting that the implementations of Modbus TCP with TLS for security were suggested in 2018 [2].

Honeypots can be used to detect new and emerging attack techniques that may not be detectable by traditional security measures. They can also provide insights into attacker behaviour and help organizations improve incident response processes [3]. Furthermore, honeypots can divert attacks from critical systems, providing additional protection. Using honeypots, organisations can identify vulnerabilities in their OT systems and take proactive measures to prevent cyber attacks that could significantly impact public safety and national security.

The honeypots are designed to simulate legitimate systems and services, comprising modules of network services such as web applications, Industrial Control Systems (ICS), and Supervisory Control And Data Acquisition (SCADA) systems. Creating an effective modular honeypot requires a comprehensive understanding of the simulated environment and regular monitoring of the honeypot's activity. Furthermore, honeypots should be isolated from production environments to prevent attackers from accessing real systems.

II. State of the art

In the article [4], the authors presented a modular honeypot system that uses containers to simulate protocols as high-interaction modules, enabling attackers to interact with the system. The system features a flexible software architecture that integrates new protocol modules. It comprises a core module for configuration and management, a packet capture and noise filter module, low-interaction modules, high-interaction modules, and an attack database.

Likewise, the article [5] outlines a proposed honeynet server architecture that utilises virtual machines and load balancing to monitor and intercept attacks on honeypots. The data is stored on a shared storage device. The honeynet uses software agents to manage and monitor traffic on the RTC monitoring cluster server and then forwards the communication to honeynet servers.

The [6] article describes a system architecture that uses containers to deploy a honeynet in a monitored network and collect data from it. The system comprises three modules: Honeynet Deployment, Intrusion Detection, and System Management. The Honeynet Deployment module uses Conpot to configure several honeypots to form a honeynet, and the Intrusion Detection module includes data collection. The System Management module enables users to manage honeypot information and intrusion detection parameters through a graphical user interface.

This article draws inspiration from previous designs by using containers for easy modularity, a load balancer for distributing the workload across servers, and shared storage for data storage. The primary focus of this article is to capture communication in the absence of preexisting honeypot services on physical or virtual networks.

III. Design

Numerous factors must be considered to ensure a honeypot system's efficacy. For the honeypots to work at their best, they must first be easy to install on any system and use few resources. Virtual Machines and Linux containers are only a few possibilities available to do this. But, Linux containers were chosen due to their simple installation on any system, scalability, and low resource needs.

Unlike Virtual Machines, Linux containers use fewer resources since they don't need a whole operating system. They are more effective and less resource-intensive since they share a kernel with the host operating system. Linux containers are the best choice for creating a modular honeypot system that is simple to build and resource-effective because of this capability [7].

The accessibility of honeypot services, which would mimic some services, is another crucial factor that needs to be considered when creating a honeypot system. The honeypot system may not operate well for some network services if no honeypots can mimic their communication. In these situations, physical or virtual services might be used as honeypots to lure attackers.

A. Honeypot system with preexisting honeypot service

The design is made of multiple parts, as shown in Figure 1. The main honeypot application comprises the Container image and the Data server. The Container image contains the Honeypot service to mimic a networking service that will communicate with the attacker.

The Base container image utilises the tcpdump to monitor all incoming or outgoing communication to the container and stores it in a pcap file. When the pcap file reaches the maximum size or the specified timeout expires, the captured communication is sent to the data server. After sending the data, the honeypot removes it and resumes listening for new communication. These parameters are intended to prevent long-term storage and ensure prompt transmission to the data server.

The vast amount of stored packets can overwhelm the security analyst. To address this, the Honeypot service can be configured to send specific data to the data server, such as attempted login usernames and passwords in SSH honeypots. Communication tools in the Base container image guarantee secure and encrypted data transmission to the data server.

The Data server is also scalable and made up of several components. The User Interface (UI) allows analysts to download and analyse captured pcap files in analyse tools, e.g. Wireshark, or obtain specific data from the honeypot service in JSON (JavaScript Object Notation) format. Additionally, the Data server includes the Application Programming Interface (API) with two endpoints for specific data and pcap files in binary format. All data is stored in a container and linked to the database or saved directly in the database, which can be located on a remote server. Lastly, the back-end component handles server logic and connects all the other parts.

B. Honeypot system without preexisting honeypot service

This section will discuss the application structure, which includes a container image and a data server, similar to the previous section. However, a container image is different because it does not contain a honeypot service. Instead, external physical or virtual devices communicate with attackers.

This situation is common with the OT protocols because they do not have many existing honeypots to mimic the communication compared to the Information Technology protocols. For this reason, the honeypot needs to monitor the

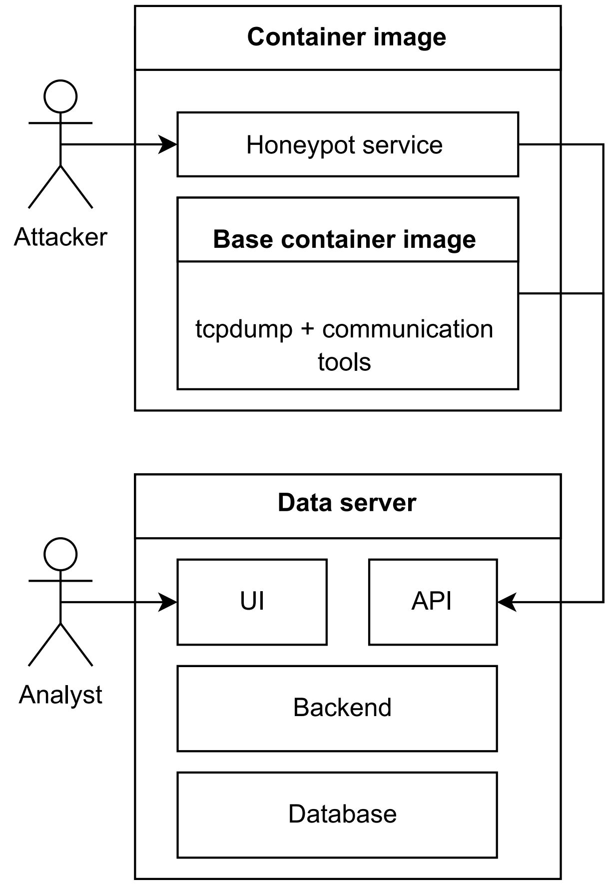


Fig. 1. Honeypot system with preexisting honeypot service design

whole communication to gather all the necessary data from the attack.

To achieve this, a network tap or switch with a Switched Port ANalyzer (SPAN) port is placed between the service and the attacker. The sole purpose of this tap or switch SPAN port is to duplicate the communication between the service and the attacker onto another port, as shown in Figure 2. By adding this interface to the container, the container can monitor and record all communication between the attacker and the service and send it to the data server.

In virtual networks, the network tap or switch with a SPAN port can also be added by creating a virtual switch, monitoring specific ports, and forwarding the communication to another port that can be attached to the virtual machine.

This approach allows security analysts to capture and analyze attackers' interactions with the service. Unlike honeypot service, this approach uses real services while capturing attacker behaviour.

It is important to note that the data server is the same as in the previous section. It has a user interface for analyzing the collected data and an API with endpoints for accessing specific data and pcap files. Data can be stored either in containers, linked to the database, or in the database itself, with the backend connecting all parts of the data server and managing the server logic.

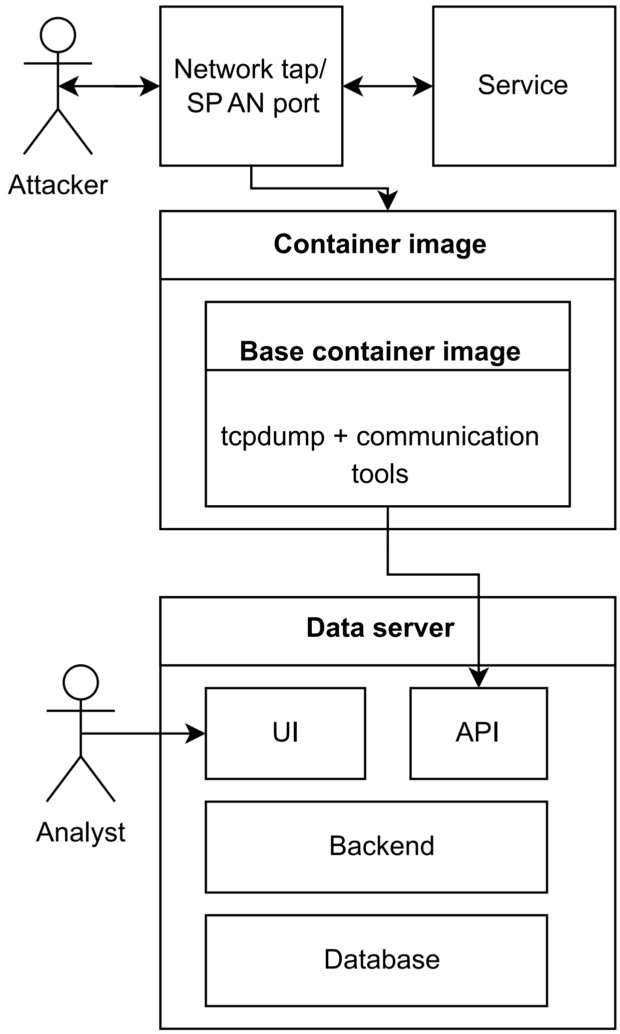


Fig. 2. Honeypot system without preexisting honeypot service design

C. Scalability

Containers are one of the most popular solutions for making applications easily installable and scalable [8]. A common practice to scale containerised applications is to have a load balancer that disperses the load between different workloads and scale-out the system by deploying more containers.

Load balancing ensures incoming requests are split evenly among several servers to maximise system performance. Diverting traffic to healthy servers in the event of a failure also helps prevent server overload and guarantees high availability.

A remote database can be added to store captured communication and data to increase the system's scalability further. As a result, each container is no longer required to maintain its database, simplifying system scaling. All data servers may direct data to the precise location with a single database, simplifying management and maintenance.

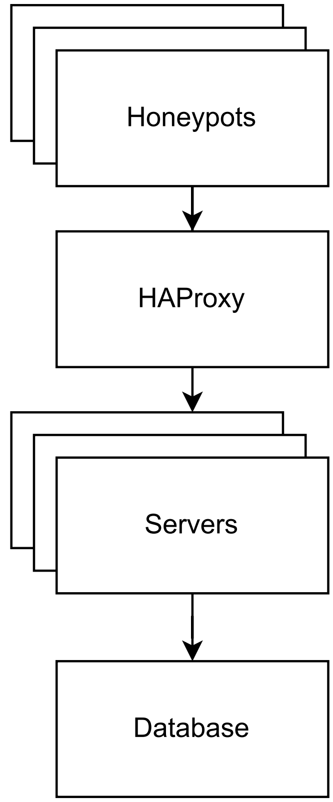


Fig. 3. Scalable design of honeypot system

IV. Implementation

Two container images were created to implement the proposed model: one for the honeypot and the other for data collection. The honeypot container was designed to simulate vulnerable systems and attract potential attackers. Meanwhile, the data collection container is responsible for collecting data on the attacks detected by the honeypot. This information is then sent to the data server container for storage and analysis.

A. Data server

The data server container was implemented using the Django Framework, well-known in the open-source community for its reliability and ease of use. One of the key advantages of using Django is that it comes with an easy-to-use model database running by default with SQLite. However, with small configuration changes, the database can be changed to any remote database running on a remote server to support the scalable design described in the previous section. The API for accessing the data collected by the data server is made using the Django Rest Framework, which simplifies the creation of RESTful APIs and provides a range of tools and features to enhance the functionality and security of the API. The Bootstrap Framework was used to create the UI for the data server to ensure that the user interface is responsive and user-friendly.

B. Container image

The container image used for the honeypot container is based on a base container image, which contains all the necessary tools for communication with the data server and monitoring the communication on the container image. The base container tools are written in bash to be as modular as possible and require only tcpdump and curl binaries. This design decision ensured the base container image could run on the most common Linux distributions, such as Fedora, Ubuntu, Debian, etc.

C. Security

The Data server creates a unique token for each honeypot to guarantee the security and integrity of the data collected from the honeypots. This token is used to authenticate and authorize communication between the data server and the honeypot. Incoming requests without a valid token will be denied by the data server, providing a message with the error. The honeypots can only add data to data servers. Even if an attacker succeeds in obtaining the token, they are not permitted to change or modify any already-existing data, but they are permitted to add new information. Although an attacker could still flood the data server with false information, the server will still hold information about how the attacker got on the honeypot.

HAProxy was added to control traffic flow between the honeypots and the data server to enhance the system's security. HAProxy allows for setting anti-DDoS attack parameters, such as limiting the number of requests per second or closing connections for attacks like slowloris [9]. HAProxy restricts the honeypots from communicating with any API endpoints other than those specified, adding a layer of security to the system.

HAProxy offers load-balancing and security features, enabling multiple servers to connect to the HAProxy. By doing this, the load is distributed evenly among the servers and the system is guaranteed to handle an increase in traffic volume [10]. The use of HAProxy enhances the system's scalability and reliability while also adding more security measures. The proposed model can offer a secure and reliable platform for gathering and analyzing attack data by considering these security considerations.

D. Virtual switch with SPAN port

To replicate the communication between an attacker and a virtual machine, it is necessary to have a virtual switch that can control the communication.

The Open Virtual Switch (OVS) was selected because of its compatibility with the most widely used enterprise virtualization platforms, including VMware vSphere and RHV. The OVS can create a virtual switch that can create numerous virtual interfaces, which can be added to Virtual Machines. These virtual interfaces can be mirrored to another virtual interface that can be added to another Virtual Machine or to a container networking namespace for monitoring the communication [11].

V. Conclusion

This article presented a design of a modular honeypot system. The system can simulate various networking protocols and includes monitoring capabilities for virtual and physical communication between services. The systems containers implementation makes it easy to use, modular and scalable design.

The future step of this design and implementation is to deploy the system onto a server to capture and analyse communication for potential security attacks. Using this honeypot system, organizations can understand the attacks against their networks and take proactive actions to prevent future breaches.

Acknowledgment

I would like to thank Ing. Petr Blažek for his expert guidance and helpful thoughts during the writing of this paper.

References

[1] Thomas Morris, Rayford Vaughn, and Yoginder Dandass. A retrofit network intrusion detection system for modbus rtu and ascii industrial control systems. In *2012 45th Hawaii International Conference on System Sciences*, pages 2338–2345, 2012.

[2] Modbus.org, "modbus/tcp security,". 2018.

[3] Iyatiti Mokube and Michele Adams. Honeypots: Concepts, approaches, and challenges. In *Proceedings of the 45th Annual Southeast Regional Conference*, ACM-SE 45, page 321–326, New York, NY, USA, 2007. Association for Computing Machinery.

[4] Shreyas Srinivasa, Jens Myrup Pedersen, and Emmanouil Vasilomanolakis. Interaction matters: A comprehensive analysis and a dataset of hybrid iot/ot honeypots. In *Proceedings of the 38th Annual Computer Security Applications Conference*, ACSAC '22, page 742–755, New York, NY, USA, 2022. Association for Computing Machinery.

[5] Ahmed Noaman, Ayman Abdel-Hamid, and Khalid Eskaf. A novel honeynet architecture using software agents. In *2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, pages 1–6, 2019.

[6] Zhiqiang Wang, Gefei Li, Yaping Chi, Jianyi Zhang, Qixu Liu, Tao Yang, and Wei Zhou. Honeynet construction based on intrusion detection. In *Proceedings of the 3rd International Conference on Computer Science and Application Engineering*, CSAE '19, New York, NY, USA, 2019. Association for Computing Machinery.

[7] Amit M Potdar, Narayan D G, Shivaraj Kengond, and Mohammed Moin Mulla. Performance evaluation of docker container and virtual machine. *Procedia Computer Science*, 171:1419–1428, 2020. Third International Conference on Computing and Network Communications (CoCoNet'19).

[8] Lily Puspa Dewi, Agustinus Noertjahyana, Henry Novianus Palit, and Kezia Yedutun. Server scalability using kubernetes. In *2019 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON)*, pages 1–4, 2019.

[9] Rizgar R. Zebari, Subhi R. M. Zeebaree, Amira Bibo Sallow, Hanan M. Shukur, Omar M. Ahmad, and Karwan Jacksi. Distributed denial of service attack mitigation using high availability proxy and network load balancing. In *2020 International Conference on Advanced Science and Engineering (ICOASE)*, pages 174–179, 2020.

[10] Yassin Abdulkarim Hamdalla Omer, Mohamed Ayman Mohammedel-Amin, and Amin Babiker A. Mustafa. Load balance in cloud computing using software defined networking. In *2020 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, pages 1–6, 2021.

[11] Seyeon Jeong, Doyoung Lee, Jian Li, and James Won-Ki Hong. Openflow-based virtual tap using open vswitch and dpdk. In *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, pages 1–9, 2018.

Soft-switching Methods and Optimization of Output Stage of MOSFET Driver

Vladimír Trojan
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
211245@vut.cz

Karel Ptáček
SCG Czech Design Center
Onsemi
Rožnov pod Radhoštěm, Czech Republic
Karel.Ptacek@onsemi.com

Roman Prokop
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
prokopr@vut.cz

***Abstract*—This paper focuses on design and comparison of optimization methods for the output stage circuit of MOSFET gate driver. Optimization methods aim to reduce voltage swing induced on parasitic series inductance of package terminals. In order to determine the effectiveness of proposed optimization methods, individual designs are compared with original MOSFET gate driver circuit designed by Onsemi company. These optimization methods should improve performance of MOSFET gate driver in the field of high-frequency use, where induced parasitic voltage can cause false signals, interference or even breakdown of unprotected structures.**

Keywords—MOSFET gate driver, soft-switching, parasitic inductance, voltage swing reduction, driver optimization, di/dt reduction

I. Introduction

Since the manufacturing of the first MOSFET in the early 1950s, this type of transistor has been implemented in almost all modern electronic devices. One of its main advantages is the ability to switch between two signal states without the need of continuous current supply, which is reflected in lower power consumption. As a result, the MOSFET has replaced formerly used bipolar transistors in many applications and is commonly used for high-frequency switching applications [1]. However, it is still necessary to supply MOSFET with sufficient current in order to charge its gate capacity. To achieve high-frequency switching, MOSFET gate drivers are implemented to supply the required current in nanoseconds [2]. Considering the real properties of the device, this may cause some issues due to the parasitic inductance of output terminals of the device package. High *di/dt* produced on output of MOSFET gate driver induces undesired voltage swing *v(t)* given by equation (1):

$$v(t) = L\frac{di(t)}{dt} \tag{1}$$

Induced voltage *v(t)* can cause false signals, interference or even breakdown of unprotected devices. This adds the limitation to maximum frequency of the application [3].

Since *v(t)* depends on *di/dt*, the solution to the problem may be implementation of soft-switching methods to reduce *di/dt* at the output of MOSFET gate driver while maintaining comparable switching frequency.

II. Optimized MOSFET gate driver

The MOSFET gate driver can be described as a power amplifier that receives low power signal input from a digital circuit and produces appropriate high current gate drive for switched power device (power MOSFET, IGBT, GaN transistors, etc.). This significantly reduces the time required to charge the gate capacitance of the power transistors and high frequency switching may be achieved, as mentioned before.

The driver, on which optimization methods were designed and tested, is one of Onsemi company design. Its output stage uses totem-pole structure, as depicted on block structure in Fig. 1. The switching time of the output PMOS and NMOS transistors and the enable state from the *EN* signal are controlled by digital circuit and are based on input signals. The digital circuit also implements dead time between the switching of output PMOS and NMOS transistors, which is crucial to avoid unnecessary power losses through cross currents. In order to isolate the digital control circuit from output power stage, it is powered by an external supply voltage V_{CC_dig}. The digital signal then has to be converted to the appropriate voltage levels to switch the output PMOS and NMOS transistors. This is achieved by using level-shifter, which transfer the digital signal on V_{CC} – *FGND* level for PMOS and V_{CCL} - *GND* level for NMOS. The voltage level *FGND* is derived from V_{CC} using a voltage regulator.

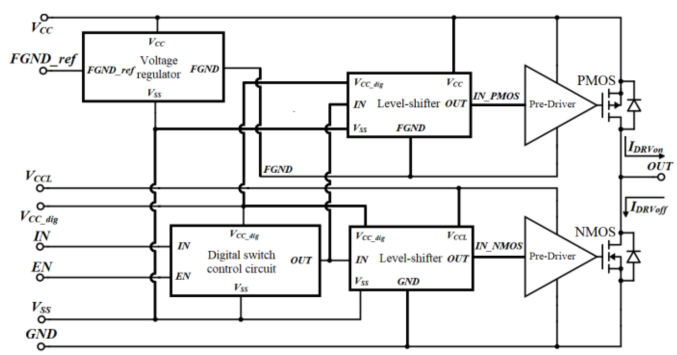


Fig. 1. Original driver block diagram

Used totem-pole structure also inverts the input signal and thus an invertor must be used to make it non-inverting structure. The output transistors also possess non-negligible gate capacity and have to be driven by set of pre-drivers. This problem is

DOI 10.13164/eeict.2023.81

solved by using inverting pre-drivers designed as cascade of invertors with increasing current capability.

III. Simulation circuit

The critical parameters of the original gate driver and later designed optimization methods were simulated and measured on a simulation circuit depicted in Fig. 2.. The simulations took place in Cadence Virtuoso environment by Spectre simulator. The gate driver with structure ,as described in Fig. 1., is supplied by PWL sources V_{1-4}. These voltage sources are connected to assumed parasitic series inductance of package terminals as they are expected to draw significant current and may influence the output parameters of the simulated gate driver. Parasitic inductance is also connected to the *GND* terminal, as majority of current from supply sources flows to it. The value of parasitic series inductance was chosen to be 3 nH as the maximal expected terminal inductance.

The input square wave signal *IN* is generated by the PWM source V_5. The *IN* input is supposed to be low power digital signal, so parasitic inductance is ignored, as well as for the PWL sources V_2 and V_6, which are also expected to be low power digital supply and enable signal. The series inductance is also neglected at the V_{SS} terminal as it is grounding potential for V_2 and V_6 sources. The series inductance is also connected to the output of the MOSFET gate driver. The capacitance C_{ISS} represents the gate capacity of switched power MOSFET connected to the output of the MOSFET gate driver.

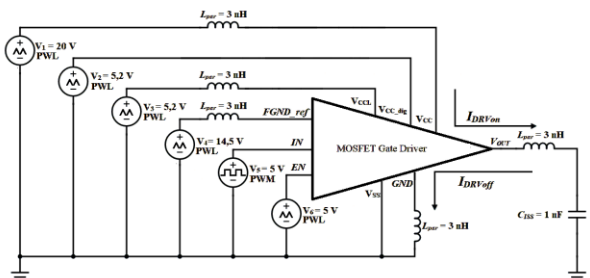


Fig. 2. Simulation circuit schematic

Parameters were measured from transient analysis in the range of 30 µs. The PWL sources were used to simulate the starting conditions of the gate driver circuit. Transient waveforms of used voltage sources and *IN* vs *OUT* signals for original driver circuit are shown in Fig. 3:

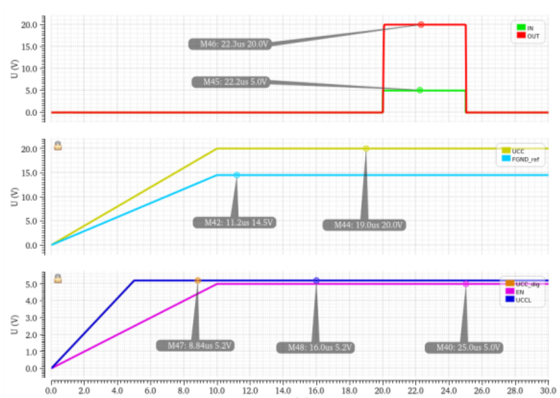


Fig. 3. Transient waveforms of V_{CC}, V_{CCL}, V_{CC_dig}, *FGND_ref, IN, EN* and *output voltage* V_{OUT}

IV. Original output stage

The output stage of MOSFET gate driver will be described in this paper as its PMOS and NMOS output transistors along with their respective pre-drivers. Original output stage circuit of MOSFET gate driver is shown in Fig. 4.. Pre-drivers for output transistors are realized as cascade of two inverters with increasing current capability. The first inverter is roughly 1/4 the size of the second inverter. This sizing ensures that the first inverter has its gate capacitance low enough to be driven by the signal from level-shifter while producing enough current to drive the gate capacitance of second inverter. The second inverter then directly controls the switching of output transistor and produces current necessary to charge its gate capacitance.

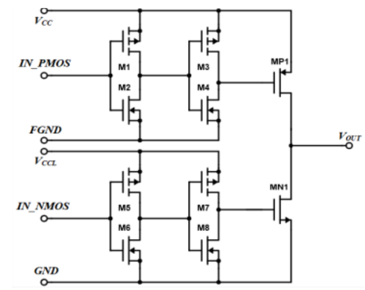


Fig. 4. Original output stage of MOSFET gate driver

The parameters measured during the simulation were its output di/dt_{rise}, propagation delay D_{rise} and rise time T_{rise}. These parameters were measured during the rising and falling edge of the driver output signal V_{OUT}. The main purpose of optimization is to reduce the voltage swing *v(t)* while maintaining the speed of driver and the measured parameters provide the best information about these attributes.

V. Method: Decreasing pre-driver size

The first method of optimizing the output stage of the original driver is to decrease the size of pre-drivers. This method does not change the wiring of the original circuit or the size of the output transistors. The only change compared to the original circuit is proportional reduction in the size of transistors in two pre-driver invertors. Decreased size of pre-driver transistors results also in decrease of the driving current for output transistors. Therefore, their gate capacitance takes longer to charge/discharge and turning ON/OFF process of output transistors is slower. Thank to that, the driving current I_{DRV} at the output of MOSFET gate driver reaches its maximum value later than in its original driver version. This increases the propagation delay, but also leads to reduction in the value of *di/dt* and therefore to reduction of voltage swing *v(t)* on its power terminals V_{CC}, *GND* and output terminal V_{OUT}. This is shown in Fig. 5, where transient waveforms of optimized driver circuit are compared to the simulation results of the original drive circuit. To compare the efficiency of each method, the optimization was designed to increase the propagation delay of circuit by only 1 ns compared to the original MOSFET gate driver circuit.

As can be seen in Fig. 5., the voltage swing *v(t)* at terminals V_{CC} and *GND* is significantly reduced, compared to the original circuit waveforms, while the propagation delay increases only by 1 ns and the maximum value of output current I_{DRV} remains unchanged. The complete results and measurements are documented in TABLE I.

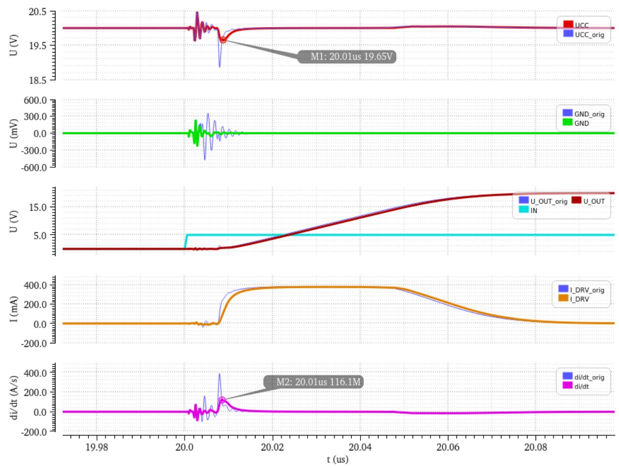


Fig. 5. Transient waveforms of optimized driver circuit compared to the original circuit (blue color) for rising edge of V_{OUT}

VI. Method: Gradual switching by delay resistors (GSDR)

The GSDR method divides output PMOS and NMOS transistors into 4 smaller segments of the total size of original transistors. The individual segments are then gradually switched by its pre-driver. The size of respective segments is also gradually scaled to 1/20, 4/20, 6/20 and 9/20 of the original transistor size. The delay between the switching of each segment is added by connecting a resistor to the gate of each segment. This is shown in Fig. 6.. The Resistor forms an RC filter with the gate capacitance of the transistor segment and the delay is given by its time constant. The resistor values are determined by simulation.

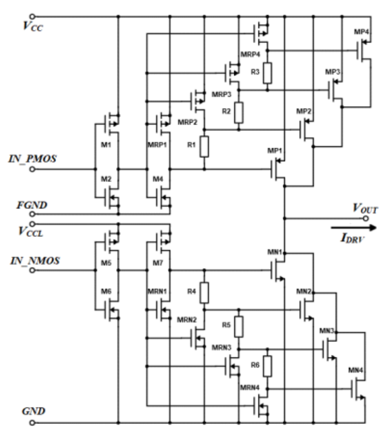


Fig. 6. Driver output stage schematic with GSDR method

The application of GSDR method results in gradual switching of individual segments of the output transistor starting with the smallest segment as the first and the largest one as the last. The resulting currents I_{DRV} on the output segments are plotted in Fig. 7.. By gradual switching of the output segments a gradual ramp-up of I_{DRV} is achieved.

The spacing between onset of current on each segment was set by the value of added resistor to achieve maximum reduction, while increasing the propagation delay again by only by 1 ns over the original driver circuit.

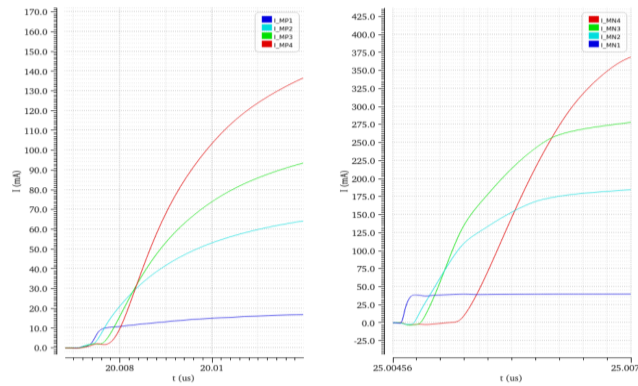


Fig. 7. Transient waveforms of I_{DRV} at output PMOS and NMOS segments using GSDR optimization method

The advantage of GSDR method is that the smallest segment, which is switched first in the row, has low gate capacitance, but it is driven by the same current as the original output transistor. As the result, the first segment turns ON faster than the original transistor and the I_{DRV} starts to rise earlier. Gained time may be used for wider current spacing. Achieved reduction of the voltage swing $v(t)$ can be seen in Fig. 8.:

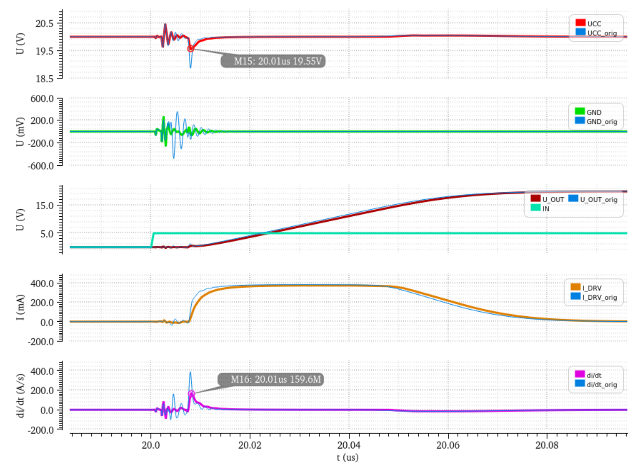


Fig. 8. Transient waveforms of driver circuit optimized by GSDR method compared to the original circuit (blue color) for rising edge of V_{OUT}

From Fig. 8., it is clear that the GSDR method achieved even greater reduction of the voltage swing $v(t)$ on V_{SS} and GND terminals, than the previous method, while maintaining the same current capability and propagation delay. The complete results for the rising and falling edges are shown in TABLE I.

VII. Method: Gradual switching by asymmetric invertors (GSAI)

Similar to the GSDR method, the GSAI method uses gradual switching of individual segments of the output transistor. However, the delay between switching of the segments is added by separate invertor for each of the segments. This is depicted in Fig. 9. The delay between switching of the output segments is controlled by size of its driving inverter. Decrease in its size results in a decrease of the drive current for its segment and the switch time of the output segment increases.

The GSAI method thus has the same advantages as the GSDR, but the delay is set much easier.

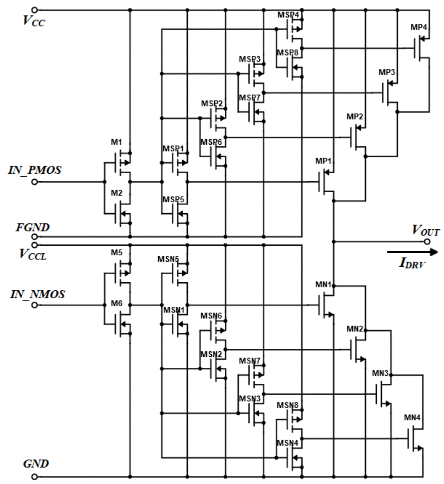


Fig.9. Driver output stage schematic with GSAI method

The resulting currents on segments of the output transistors are shown in Fig. 10., where the spacing between onset of current is more precise than in the GSDR method.

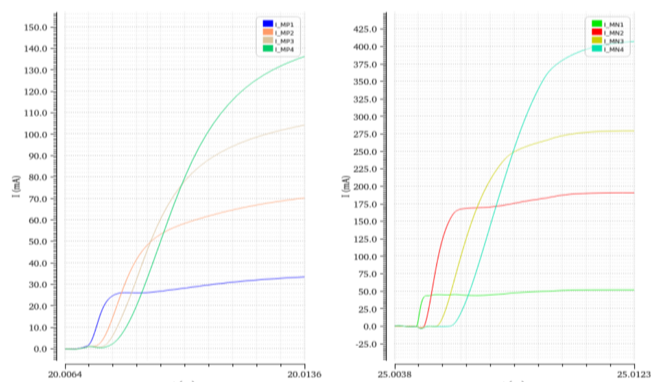


Fig. 10. Transient waveforms of I_{DRV} at output PMOS and NMOS segments using GSAI optimization method

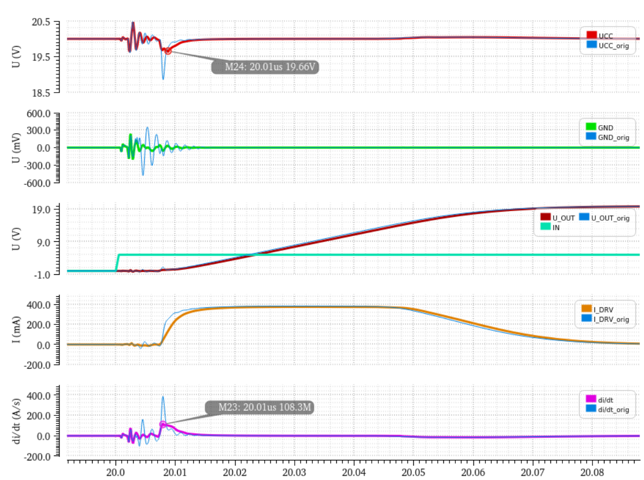


Fig. 11. Transient waveforms of driver circuit optimized by GSAI method compared to the original circuit (blue color) for rising edge of V_{OUT}

As with the GSDR method, this type of switching starts onset of I_{DRV} current earlier than the original driver output stage and the gained time may be utilized for *di/dt* reduction. The GSAI method is compared with the original driver waveforms in Fig. 11.. According to these simulation results, the GSAI method has achieved the best voltage swing *v(t)* reduction, while maintaining 1 ns additional propagation delay and maximal I_{DRV} value. Complete measurement results are documented in TABLE I.

TABLE I. SIMULATED PARAMETERS OF ORIGINAL CIRCUIT AND OPTIMIZATION METHODS

D_{rise} [ns]	D_{fall} [ns]	T_{rise} [ns]	T_{fall} [ns]	di/dt_{rise} [A/ns]	di/dt_{fall} [A/ns]
Original driver parameters					
13,97	7,19	44,59	18,51	0,278	0,709
Method: Decreasing pre-driver size					
15,07	8,21	45,58	19,29	0,097	0,230
Method: Gradual switching by delay resistors (GSDR)					
15,26	8,27	45,44	19,35	0,104	0,333
Method: Gradual switching by asymmetric invertors (GSAI)					
15,15	8,20	45,68	19,60	0,106	0,261

VIII. CONCLUSION

In this paper, the means to reduce the voltage swing *v(t)* induced on the parasitic inductances of MOSFET gate driver terminals were presented. This was achieved by reducing *di/dt* at the output of MOSFET gate driver. Three methods have been presented. The first method is the simplest to implement, as it does not change the wiring of the original driver output stage. However, this method does not perform as good as other two methods. The GSDR and GSAI methods share the same advantages, but the GSAI method has achieved better performance through more precise spacing between onsets of current on the output transistor segments. All proposed optimizing methods have reduced *di/dt* on the output terminal of MOSFET gate driver almost three times at the cost of only 1 ns of additional propagation delay. These methods have been simulated on a real commercially used circuit of Onsemi company. Furthermore, in the final version of the diploma thesis, the proposed circuits will be manufactured and connected to SOIC-16 packages. Manufactured circuit will be tested in real conditions and real parameters of used methods will be determined.

REFERENCES

[1] L. Balogh, „Fundamentals of MOSFET and IGBT Gate Driver Circuits,“ Bedford, New Hampshire, USA, 2018.

[2] Sipex Corporation, Properly Sizing MOSFETs for PWM Controllers,“ USA, 2006.

[3] B. K. Bose, Modern power electronics and AC drives, Upper Saddle River, NJ: Prentice Hall, 2002.

Preparation of Inversely Vulcanized Sulfur for Use in Post Lithium-ion Batteries

David Trochta
Department of Electrical and Electronic Technology
Brno University of Technology
Brno, Czech Republic
199998@vut.cz

Ondřej Čech
Department of Electrical and Electronic Technology
Brno University of Technology
Brno, Czech Republic
cechondrej@vut.cz

***Abstract*—This paper focuses on the synthesis of various types of inversely vulcanized sulfur samples, the description of their properties, and their structural analysis with X-ray diffraction and Raman spectroscopy. Then, the process of manufacturing electrodes with inverse vulcanized sulfur is described. The electrochemical properties are then tested by galvanostatic cycling with a potential limitation in lithium-sulfur battery.**

Keywords*—*inverse vulcanization, inversely vulcanized sulfur, Li-S battery, post lithium-ion batteries

I. Introduction

Lithium ion (Li-ion) batteries are currently the most popular type of batteries for portable devices or electromobility. However, requirements for these applications are increasing, and Li-ion batteries cannot provide necessary properties such as capacity or a low cost of the used materials [1].

Due to this, research is being done on alternative battery storage systems that are intended to replace Li-ion batteries at least partially in various applications. These new types of batteries are generally called "post lithium-ion batteries", and they also include lithium sulfur (Li-S) and room temperature sodium sulfur (RT Na-S) batteries. These batteries are using cheap and nontoxic sulfur as electroactive material, which also results in high on paper energy density, which goes up to 2600 Wh/kg for Li-S [2] and 1230 Wh/kg for RT Na-S [3]. These values exceed the theoretical capacity of the Li-ion, which goes up to 600 Wh/kg [1]. On the other side, there are still some unresolved issues that cause the low longevity of the Li-S and RT Na-S batteries due to the use of sulfur [2].

Various studies are currently being conducted on how to increase the lifespan of these systems [2]. One of the possible approaches is use of sulfur in amorphous form, which can be obtained by using inverse vulcanization. However, this process can be carried out in several possible ways, which leads to different types of inversely vulcanized sulfur (IVS) with different mechanical and electrochemical properties. The most important property, however, is the ability to resist reverse recrystallization and thus maintain its amorphous structure in the long term [4].

II. Inverse Vulcanization

Vulcanization is a generally known process for producing or hardening rubbers. In this process, sulfur forms bridges between sections of the polymer chains and thus changes the properties of the resulting product. But in 2013 Griebel, Chung et al. presented a method called inverse vulcanization and at the same time showed a final product called IVS and its properties with possible applications [4].

The whole process of the inverse vulcanization consists in exchanging the roles of the precursors. Sulfur no longer forms bridges between polymer chains but forms polymer chains itself. These sulfur chains are subsequently interconnected (cross-linked) by some other monomer (crosslinker). In the original work from 2013, the authors decided to test and use 1,3-divinylbenzene (DIB) as the crosslinker [4].

Using their procedure, the authors succeeded in synthesizing several IVS samples with different ratios of input precursors. They tried ratios (S:DIB) from 95:5 wt% up to 50:50 wt%, but only the samples with 20 wt% of DIB and more were found to be amorphous and chemically stable. Subsequently, the authors also verified the use of IVS as electroactive materials in Li-S batteries, thereby confirming that the research on this material has a further meaning not only in electrochemical applications [4]. To this day, the IVS has also been tested in applications like thermal insulators [5], mercury absorbents [6], fertilizers [7], and many more.

III. Methods of IVS Preparation

We developed and optimized several procedures for preparing IVS, which differ mainly in their length or maximal temperature during the synthesis. For the preparation of the IVS samples in this paper elemental sulfur from Sigma Aldrich with a purity >99,5% was used. As the crosslinker, DIB from TCI with 97% purity was used. The remaining 3% is 4-tert-butylcatechol (TBC) stabilizer, which prevents polymerization of the DIB itself during transportation and storage. This stabilizer was partially removed right before a synthesis by stirring DIB with a 4 M aqueous solution of sodium hydroxide (NaOH).

DOI 10.13164/eeict.2023.85

The synthesis apparatus itself is essentially an oil bath with the possibility of temperature regulation and mixing using a magnetic stirrer. A test tube is then inserted into this oil bath, in which the IVS synthesis itself takes place. In addition, to the necessary precursors (sulfur and DIB), there is also a magnetic stirrer in the test tube, which ensures the stirring of the mixture during the synthesis.

A. High Temperature Synthesis

This process, which can also be seen in Fig. 1. (b), started with the heating of elemental sulfur. During the heating, sulfur melted (around 125 °C) and turned into the yellow liquid, which slowly turned into the orange colour. When the mixture reached 155 °C, a small amount of DIB was added to ensure future miscibility. The orange liquid slowly turned transparent red, and when the temperature reached 170 °C an aliquot amount of DIB was added to reach the required final S:DIB ratio. Melted sulfur and DIB formed two layers, but after a short period they mixed into a homogenous mixture. After that the temperature was set to 185 °C. The whole mixture slowly changed its colour from a transparent orange to a non-transparent dark red. During the colour change, the mixture also increased its viscosity, and the magnetic stirrer stopped rotating. Then the tube was taken out of the oil bath and moved to an oven that was preheated to 140°C for final curing, which took 8 hours. After the curing, IVS was cooled to room temperature. The final product is called a marked as IVS-HT (S:DIB ratio in wt%). With this method, we were able to synthetize various IVS-HT samples with different S:DIB ratios, but only samples with 30 wt% DIB and more were stable and suitable for further research.

After removing IVS-HT samples from the tubes by breaking them with a hammer, chunks of IVS-HT were obtained. These IVS-HT chunks can be seen in Fig. 1. (c). After that, elementary properties of the samples could be observed with the naked eye. The characteristic properties for this IVS type are elasticity, stickiness, and low resistance to higher temperatures. For the last reason, these samples must be stored in a freezer. It is also important to note that the individual samples with different S:DIB ratios do not differ in color or other elementary properties that can be investigated by the naked eye.

B. Low Temperature Synthesis

The problem with the high temperature method is that DIB is evaporating during it, and for that reason the final ratio of the precursors is not certain. Therefore, another method for IVS preparation was developed and optimized. The main advantage of this method is a lower maximum temperature during the inverse vulcanization, which makes DIB evaporate less and the final precursor ratio more accurate. The lower temperature process is also easier and less energy consuming.

The lower temperature process, which can be seen in Fig. 1. (b), started the same way as the high temperature one. The sulfur was heated, it melted, and the temperature was set at 155 °C. At this temperature, the whole amount of aliquot DIB was added, and the temperature was set to 165 °C. As in the previous method, sulfur and DIB formed two layers, which mixed after a while. This time the transparent orange mixture changed the colour to transparent red. During this change, the viscosity also increased, and thus the magnetic stirrer stopped. The curing in the preheated oven followed at temperature of 110 °C for 8 hours. After the curing the IVS was cooled to a room temperature. The final product is called and marked as IVS-LT (S:DIB ratio). With this procedure we, were able to synthesize stable samples with 20 wt% DIB and more. The final products were also taken out of the tubes by breaking them. The obtained chunks of IVS-LT can be seen in Fig. 1. (d).

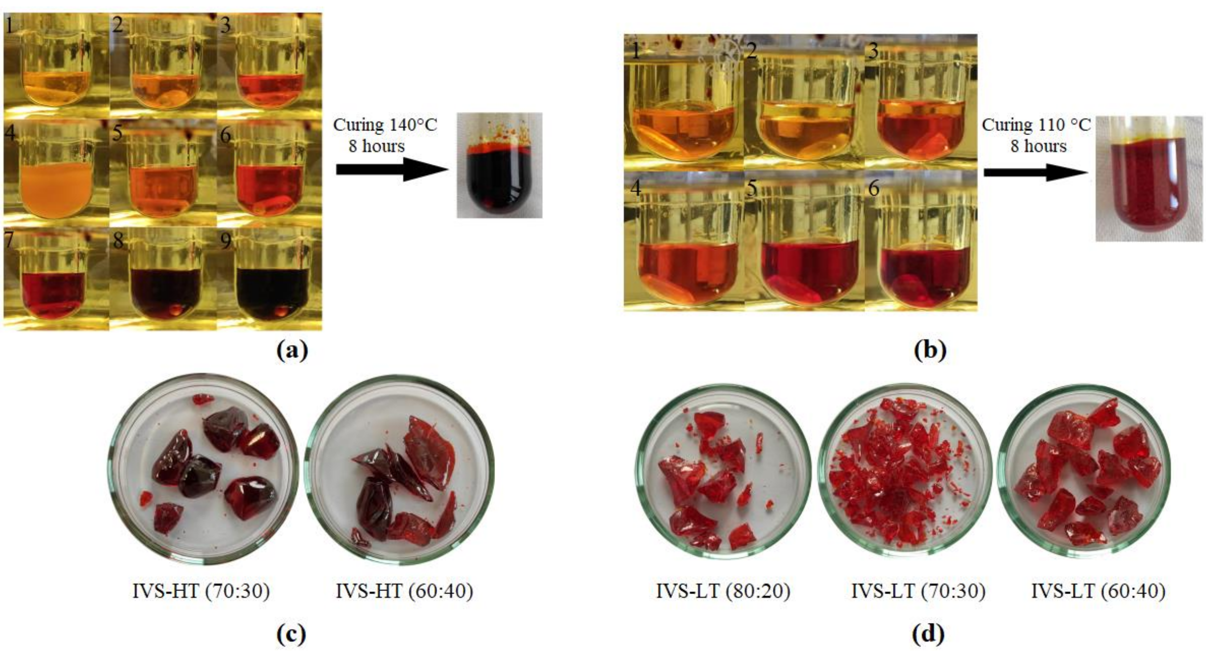


Fig. 1. (a) The high temperature synthesis of IVS (IVS-HT) (b) The low temperature synthesis of IVS (IVS-LT) (c) Final products IVS-HT with a different ratios (d) Final products IVS-LT with a different ratios

As in the case of IVS-HT, individual samples of IVS-LT with different S:DIB ratios are indistinguishable from each other with the naked eye. However, other properties are different. The IVS-LT chunks are solid, hard, and non-sticky. Thus, they can be stored at room temperature, and they are much easier to process.

IV. Structural Measurements

After the synthesis, the samples were left in the test tube for several days and were observed for any signs of recrystallization. If there was no visible recrystallization, the amorphous structure of the samples was verified using the XRD and Raman spectroscopy. Thanks to these methods, we can be sure that the IVS samples are truly amorphous and suitable for further electrochemical investigation in Li-S or RT Na-S batteries.

A. *X-ray Diffraction*

X-ray diffraction (XRD) is primarily designed and used for the investigation and determination of crystal structures. Since elemental sulfur is crystalline, it can be easily detected by XRD, and it will show up in a diffractogram as characteristic peaks. Therefore, to confirm the amorphous structure of the IVS samples, there must not be any characteristic peaks of elemental sulfur in the measured diffractogram. Otherwise, if any characteristic peaks are presented, it means that there are unreacted residues of elemental sulfur in the IVS samples that will act as nuclei for further recrystallization.

The normalized diffractogram of the IVS samples, along with the diffractogram of elemental sulfur, can be seen in Fig. 2. (a). It was measured in a range from 5° up to 60° with a scan speed of 10°/min. From the measured results, it can be seen, that all prepared and presented samples are amorphous, as there are no characteristic elemental sulfur peaks in them. The various small visible peaks are just background noise, which can be caused by many factors, e.g., gamma radiation or a sample holder.

B. *Raman Spectroscopy*

Since the amorphous structure was confirmed by XRD, using Raman spectroscopy may seem redundant, but that is not true. With Raman spectroscopy, in addition to the structure and its amorphous or crystalline state, we can also examine the quantitative ratios of the used precursors. The problem is that the Raman signal and its intensity depends not only on the amount of substance in the material but also on its polarizability. Due to this, we cannot measure and calculate the exact final ration of sulfur and DIB so far, but we can only compare different IVS samples to each other. This is, for example, very convenient in a situation where it is necessary to compare different batches that were prepared in the same way and with the same S:DIB ratio, but also for comparing different IVS samples with different rations.

The last-mentioned use is presented in Fig. 2 (b). Raman IVS-HT spectra were measured by laser with 532 nm wavelength, with 2 mW power and then normalized. Low power is important, otherwise, the samples degrade, and the measured spectra are distorted. In the elemental sulfur spectrum, there are two peaks between 300 cm^{-1} and 600 cm^{-1} Raman shift. By observing these two peaks, it can be determined that the structure is crystalline. In the IVS-HT spectra, these two peaks are merged, which signals amorphous structure. This new peak is always the highest one in the IVS-HT spectra. There is also another important peak, which is highlighted by dashed lines. This peak, at approximately 1000 cm^{-1} belongs to DIB. In the original DIB spectrum, it is not the highest one, but it is more suitable for our purposes because this peak changes significantly depending on the S:DIB ratio. As expected, the peak has the highest intensity in the IVS-HT sample with a 50:50 wt% ratio, and the lowest intensity is in the IVS-HT sample with a 70:30 wt% ratio. There are also other DIB peaks that follow this trend. More precisely, they are between 1400 cm^{-1} up to 1600 cm^{-1} and 2800 $cm^{-1}$and 3100 cm^{-1}, but these are not so easily observable.

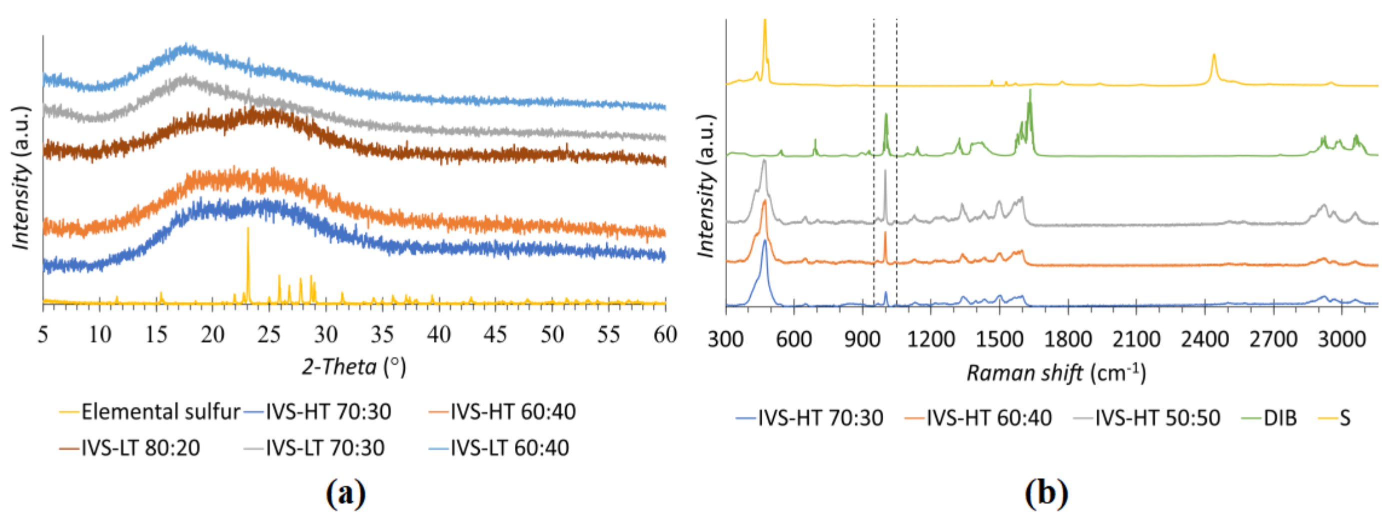


Fig. 2. (a) Diffractogram of the IVS samples, along with elemental sulfur one for comparison (b) Raman spectra of the IVS-HT samples, sulfur and DIB

V. Electrochemical Measurements

To investigate the electrochemical properties such as capacity, IVS-HT 70:30 and IVS-LT 80:20 were selected as representatives of each presented synthesis procedure with the highest sulfur content. After that, electrodes were made from both selected samples.

A. Electrodes Manufacturing

Before an electrode slurry can be prepared, it is necessary to grind IVS chunks into a powder. A vibration mill was used for this purpose, with IVS being placed inside for 10 minutes. To increase IVS's brittleness, the mill container and IVS were placed in the freezer prior to grinding.

This IVS powder was mixed with Super P carbon and carboxymethylcellulose (CMC), which is a binder, in a ratio of 60:30:10 wt%. As a dissolvent for CMC, a mixture of demineralized water and ethanol in a 2:1 v. ratio was used. The whole mixture was first stirred by hand, and then five 10 mm diameter grinding balls were added, and the mixture was moved into a planetary mill. The planetary mill was set to 350 rpm and 20 cycles. Each cycle contained 20 minutes of mixing and a 10-minute break to allow the mixture to cool.

Afterwards, a 200 μm layer of the electrode slurry was coated onto aluminium foil using a coater and left to dry. The drying process started out at room temperature and then moved on to a dryer set at 60 °C. The drying was followed by cutting 18 mm diameter electrodes and their pressing using a force of 1600 kg/cm^2.

B. Cell Assembly

The assembly of the cell within a glovebox with an argon atmosphere is the final phase. The metal lithium anode, manufactured cathode, and separator soaked with an electrolyte were put together and hermetically sealed. The used electrolyte was 0,7 M lithium bis(trifluoromethanesulfonyl)-imide (LiTFSI) + 0,3 M lithium nitrate ($LiNO_3$) in a solution of dioxolane (DOL) and dimethoxyethane (DME) in a 2:1 v. ratio.

C. Galvanostatic Cycling with Potential Limitation (GCPL)

The GCPL was performed in a potential window from 1,8 V to 2,6 V. In total, 30 cycles were performed with a 0,2C current. The results of both electrodes (Li-S cells) can be seen in the summary Fig. 3. The initial capacity of the IVS-HT (70:30) sample was 660 mAh/g (mass of sulfur contained), and after 30 cycles, it decreased to 612 mAh/g, which means a 7% capacity loss. The second electrode with IVS-LT 80:20 had an initial capacity of 422 mAh/g and decreased to 375 mAh/g after 30 cycles, which means 11% capacity loss.

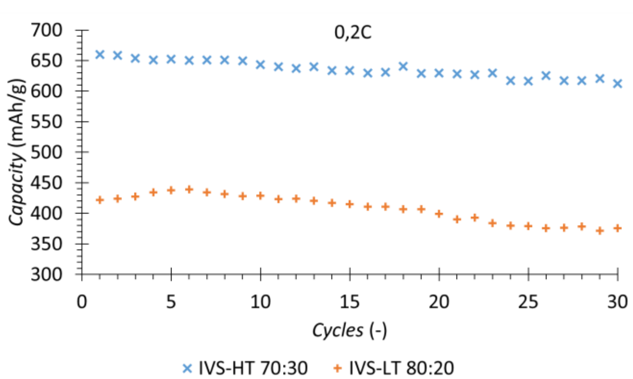


Fig. 3. Cycling performance of the selected IVS samples

VI. Conclusion

In this paper, the method of obtaining amorphous sulfur by inverse vulcanization was introduced and used for the synthesis of inverse vulcanized sulfur. Afterwards, its application in post lithium systems was presented, and other applications were also mentioned. Two different methods (high temperature and low temperature) of IVS preparation with 1,3-diizopropenylbenzene as the crosslinker were also presented. The properties and differences of the final products of both methods were investigated and described. In addition, they were measured using X-ray diffraction to confirm their amorphous structure and with Raman spectroscopy to investigate the final S:DIB ratio. From the synthesized samples, those with the highest sulfur content were then selected, and electrodes were made from them. Li-S cells were then assembled from these electrodes and tested using galvanostatic cycling with potential limitation.

The electrode with the electroactive material IVS-HT 70:30 achieved better results because it had a higher initial capacity (660 mAh/g) and lost 7% of its initial capacity after 30 cycles, which is less than in the case of the second tested electrode with IVS-LT 80:20, which had an initial capacity of 422 mAh/g and its loss after 30 cycles was 11%.

Through this approach, we have proven the viability of different IVS-based electrodes and their potential for further study and optimization for use in the battery industry.

References

[1] Y. Miao, L. Liu, Y. Zhang, Q. Tan and J. Li, "An overview of global power lithium-ion batteries and associated critical metal recycling", *Journal of Hazardous Materials*, vol. 425, 2022.

[2] T. Liu, H. Hu, X. Ding, H. Yuan, C. Jin, J. Nai, Y. Liu, Y. Wang, Y. Wan and X. Tao, "12 years roadmap of the sulfur cathode for lithium sulfur batteries (2009–2020)", *Energy Storage Materials*, vol. 30, pp. 346-366, 2020.

[3] D. Kumar, D. Kanchan, S. Kumar and K. Mishra, "Recent trends on tailoring cathodes for room-temperature Na-S batteries", *Materials Science for Energy Technologies*, vol. 2, no. 1, pp. 117-129, 2019.

[4] W. Chung, J. Griebel, E. Kim, H. Yoon, A. Simmonds, H. Ji, P. Dirlam, R. Glass, J. Wie, N. Nguyen, B. Guralnick, J. Park, Á. Somogyi, P. Theato, M. Mackay, Y. Sung, K. Char and J. Pyun, "The use of elemental sulfur as an alternative feedstock for polymeric materials", *Nature Chemistry*, vol. 5, no. 6, pp. 518-524, 2013.

[5] R. Parreño Jr, Y. Liu and A. Beltran, "Sulfur copolymers (SDIB) from inverse vulcanization of elemental sulfur (S_8) for polymer blend", *IOP Conference Series: Materials Science and Engineering*, vol. 778, no. 1, 2020.

[6] J. Chalker, M. Mann, M. Worthington and L. Esdaile, "Polymers Made by Inverse Vulcanization for Use as Mercury Sorbents", *Organic Materials*, vol. 03, no. 02, pp. 362-373, 2021.

[7] S. Valle, A. Giroto, R. Klaic, G. Guimarães and C. Ribeiro, "Sulfur fertilizer based on inverse vulcanization process with soybean oil", *Polymer Degradation and Stability*, vol. 162, pp. 102-105, 2019.

Optical fibres for ionising radiation measurements

Tadeáš Zbožínek
Department of Electrical and Electronic Technology, FEEC
Brno University of Technology
Brno, Czech Republic
214774@vut.cz

***Abstract*—The work focuses on studying methods of measuring ionizing radiation using optical fibres. More attention is paid to their use, together with scintillation materials. The proposed method allows the connection of an electronic detector with a scintillation sensor by optical means. The interconnection provides the possibility of shielding the detector from the effects of high ionizing radiation activities. The shielded detector will thus not be subject to the adverse effects of this radiation and can be used even for long-term measurements of high activities. First experimental measurements of reference sources verify the principle of the proposed method.**

Keywords—Ionizing radiation, scintillation detectors, ionizing radiation measurement, optical fibres.

I. Introduction

Ionizing radiation has been utilized widely for a very long time in many different fields. A significant increase in detection possibilities has resulted from its widespread use, which is crucial for the preservation of human life. Geiger-Müller counters and photosensitive materials were used in the beginning. Further options, including the utilization of semiconductor structures or the scintillation of specific materials, were gradually introduced. All these techniques can be used to measure ionizing radiation activity. However, exposure to high activities usually damages the detection electronics and shortens the detector´s lifetime.

To get around this issue, optical fibres are used. This prevents ionizing radiation from harming the typically expensive detector electronics. Yet, the use of optical fibre as a radiation sensor also has drawbacks due to its limited range, low sensitivity, and typical requirement for material implementation.

The technique we suggest combines the utilization of optical fibres with a scintillation crystal-based luminous sensor. This will eliminate the requirement for optical fibres that require special modifications and open the way for optical fibres that are highly resilient to radiation damage and commercially available. Moreover, using a scintillation crystal enhances the measurement range and adds the required sensitivity. The overall design will provide a way to position the electronics away from radiation damage. As a result, we can measure high gamma radiation doses over long periods of time [1] [2].

II. Method description

There are three major elements to the entire measurement. They are the detector, communication path, and sensor. The detector is positioned outside of radiation range. For our measurement, we are using primarily photomultiplier tubes (PMTs). Another detector, such as a silicon photomultiplier or single photon counter, might be used as an alternative. The communication path consists of optical fibre with high radiation resistance. Optical fibre servers purely as transport for photons between sensor and detecting PMT. The final part is scintillation material, which is exposed to radiation. We are using inorganic materials, also called scintillation crystals.

A. Principle

Radiation interacts with scintillation crystal. Thanks to material properties, specifically fluorescence, material is able to convert high-energy photons into ones with lower energy. Newly formed photons have wavelengths that fall inside the visible-light spectrum. One photon of gamma radiation creates several visible photons, resulting in the creation of a light pulse. Thus, light pulses hold information about activity and energy. A schematic connection is displayed in Fig.1.

Light pulses are routed to optical fibre. Pulses are able to travel long distances due to low attenuation of optical fibres. By doing so, we can easily lead the light signal from the measured location and protect the detector from radiation exposure. The light pulse is then caught by detector and converted to an electrical pulse for later calculations. Detection

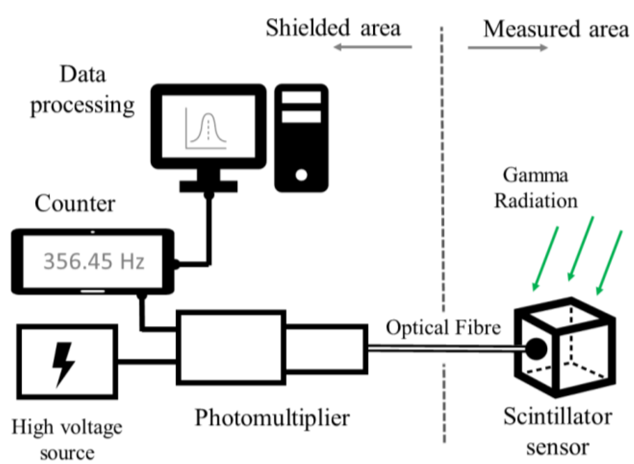


Fig. 1. Scheme of the proposed method

depends on device used. PMT maintain information about energy in form of amplitude and activity in form of pulse count. After signal processing, there are many things we can learn about the initial radiation, including, for example, which particle caused the pulse. For the purpose of this work, the activity of gamma radiation is measured. Then, one of the previously mentioned detectors can be used. Especially single photon counter which only provides information in the form of pulse count [2].

B. *Scintillation crystal*

We are using scintillation materials for the conversion of high-energy particles (in our case, high-energy photons of gamma radiation) to visible light. This process is possible with some materials thanks to fluorescence. Fluorescence is a very quick reaction and creates light photons almost immediately after radiation passes through material. Creating a thin light pulse to observe. There are other processes that are competing with fluorescence and decreasing the effectiveness of the scintillator. Especially phosphorescence and slow fluorescence which create after-glow effect. Last mentioned is loss due to heat transfer in material [1].

There are two types of scintillation materials: inorganic and organic scintillators. Each material has a different mechanism for luminescence therefore each has unique advantages. It is high light-yield for inorganic materials and a quick response for organic ones. We choose to use cerium doped lutetium-based scintillation crystal (LYSO). It is inorganic material and has one of the highest densities among scintillators. Materials with higher densities are better at absorbing radiation. Also, LYSO is one of the materials with a higher light yield. These are reasons for choosing this type of scintillators [3] [4].

The scintillation crystal is stored in a steel case that prevents random light interference. When radiation passes through crystal, it creates photons with random directions. Random direction creates a great loss in signal strength because optical fibre is attached only to one side. To reduce this loss and increase signal output, we cover the crystal with Teflon tape. Teflon works perfectly as a mirror for light waves generated by LYSO crystal. This way, we can link more photons to fibre and reduce light loss [2]. The LYSO crystal is shown in Fig. 2.

C. *Optical fibre*

We use multimode optical fibres with higher radiation resistance covered with steel protection tube to extend the life of optical fibres and avoid mechanical damage. We achieve resistance to radiation by choosing a pure silica core of optical fibre. A larger core diameter and a high numerical aperture (NA) are also essential properties. By doing so, we may extend the distance between the sensor and the detector, the life of the fibre, and the quality of light signal extraction on the sensor side [5] [6].

Multimode optical fibres are obtained from Thorlabs with the marking FP1000URT and FP1500URT. The number marks the core diameter in micrometres. Refractive index of fibres is 1,458434 and NA is 0,50. Higher hydroxyl content of these fibres decreases attenuation in the UV parts of the transmitted spectra [7].This is necessary because scintillator we are using

Fig. 2. Steel case (left) for protecting scintillation crystal (right)

has a light emission peak around 415 nm [3]. This modification decreases signal loss but also decreases radiation resistance. Optical fibres need to be adjusted, properly shortened, and shielded with protection to make measurements repeatable. This process is described in Chapter 3.

D. *Detection*

The second end of the optical fibre is attached to the detector. Light pulses have a very low level of amplitude. To be properly read, we are using a photomultiplier tube. The photomultiplier tube converts light pulses into an electrical signal. Every incident photon reacts with the electrode by emitting an electron. The electron then travels via tube and hits a system of dynodes. Every electron that hits a dynode emits extra electrons as a result of secondary emission. This way, we can get an electric pulse of higher amplitude, which can be much easier to process. Mechanism of signal conversion enables to keep information about amplitude of former light pulse. Amplitude provides details about radiation energy. As mentioned, we are focusing on radiation activity. The activity of the source can be calculated from the number of detected pulses. Pulses can also be detected by the single-photon counter (SPC). The SPC principle is similar to that of an avalanche diode. An incident photon launches an electron avalanche that creates an electrical pulse. The SPC has the advantage of not requiring a high voltage source [1] [8].

III. OPTICAL FIBRES PREPERATION

One of the main procedures during the preparation for the measurement of gamma radiation with the proposed method is the preparation of optical fibres. Measuring needs to be repeatable, adjustable for fibres of various lengths, and resistant to noise from ambient light. To secure the same distance between the sensor and optical fibre, we are using SMA905 connectors. Optical fibre is cut and cleaved for the required length. During this process, parts of the plastic coating near the ends of fibres are removed. The removal of the coating allows the adhesion of connectors. Before adding connectors, the fibre is first shielded with a black shrink film that protects it from ambient light. The next protection layer is stainless steel tube, which gives fibre better mechanical properties. Fibre with a protection is ready for connector gluing. After the adhesive has dried, the procedure continues with the grinding and polishing of the fibre ends. The ends of the optical fibres are aligned with the connectors by grinding. Grinding and polishing also remove grooves from fibre surfaces. This way, we decrease signal loss on both ends due to random light refraction. The result of polishing is shown in Fig. 3. Fibres are polished with diamond lapping sheets. Sheets have diamond grit sizes of 30, 6, 3, and 1 μm. Diamond grit with a size of 0.02 m is used

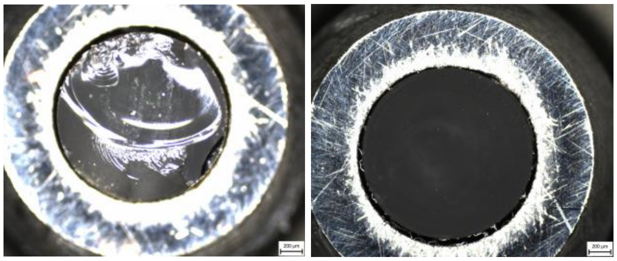


Fig. 3. Zoom on cleaved optical fibre (left) and polished optical fibre (right) [2]

for the final (optional) polishing. The whole polishing process requires cleaning with isopropyl alcohol and compressed air. Because the core diameter is larger than usual, periodic cleaning and inspection under an optical microscope are needed. Small, uncleaned particles can scratch the fibre surface, and the whole polishing process must be repeated. Fibre is now ready for the measurement. Several fibres are prepared by this procedure [2].

IV. Activity measurement

We are using reference sources for measuring the impact of changes in the length and diameter of optical fibres. These are isotopes cobalt 60 and caesium 137 with different levels of activity. Activity is measured in Becquerels (Bq), where 1 Bq equals 1 disintegration per second. We compare the number of measured pulses (counts) to the value of the reference source. It is essential to keep in mind that while cobalt and caesium have similar activities, cobalt produces more gamma rays than caesium due to distinct decay schemes. Beta rays are also produced during the decay process but are shielded by the casing of the source and scintillator. Thus, creating two types of gamma-radiant etalon with three different activities for each type [1].

A. Measurement

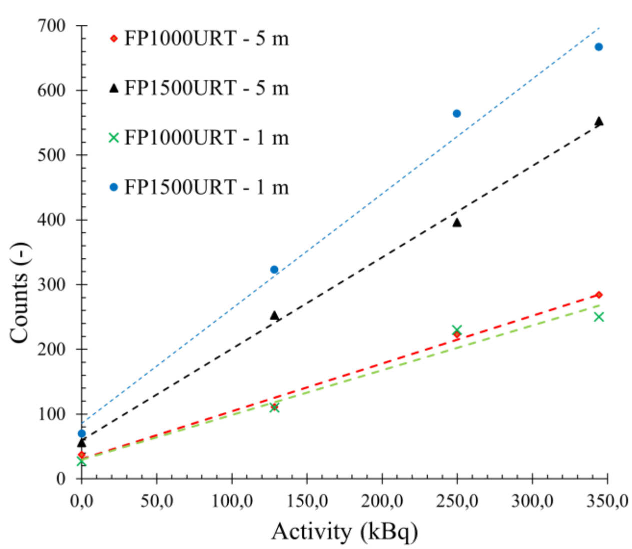
Prepared optical fibres are connected via connectors to PMT and scintillation crystal. A black box is then used to store the measurement setup, and black tape is used to seal the holes for the connectors. This will lessen ambient light noise. Then, measurements are made by laying a reference source close to the scintillator. Pulses from PMT are counted every second for ten minutes. After measuring ends, an optical fibre is replaced by another, and the process repeats. Results from measurement are shown in Fig. 4 for both reference sources.

B. Results

Four optical fibres were measured. Two optical fibres with different diameters and two optical fibres with different lengths. Graphs in Fig. 4 show that for fibres with 1 mm diameter, there is a very small change in counts between different lengths. The larger difference is between optical fibres with different diameters. Greater surface area allows more photons to enter the optical fibre. Every fibre has also almost linear progression. This is an actual advantage compared to other measuring methods. Some methods require another material implementation to secure their linearity. This often decreases their resistance to radiation damage [1] [5]. The difference between FP1000URT – 1m and FP1000URT – 5 m is almost none. In Fig.4 is shown that longer optical fibre has lower signal loss. This is due to the low attenuation of silica fibre, and the difference could be caused by the polishing process. This is necessary for polishing optical fibre. Even small portions of scratches can cause differences in the final number of counts.

V. Conclusion

We have presented a new method for long-term measurements of gamma radiation. The method is based on optical fibres and the use of a scintillation crystal as a sensor. Thanks to this setup, we are able to measure continuously. Furthermore, as demonstrated by the results, a greater distance from the sensor can be achieved. Allowing us to place electronic



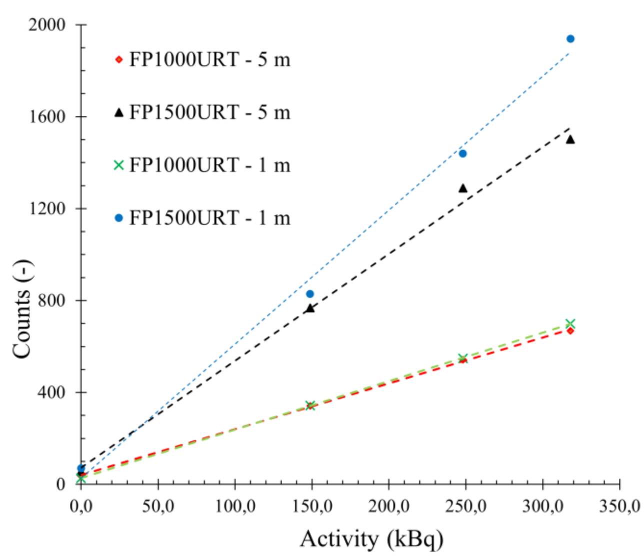


Fig. 4. Activity measurement for two different reference sources: caesium 137 (left) and cobalt 60 (right)

devices at further distances and reducing the need for human actions in measured areas. Results have also shown good linearity thanks to the scintillation crystal. Thus, enabling easier activity determination.

ACKNOWLEDGMENT

Many thanks to Ing. Břetislav Mikel Ph.D., and Ing. Jelínek Michal for their technical discussions and assistance with setting up the experiment.

REFERENCES

[1] G. Knoll, *Radiation detection and measurement*, 4th ed. Hoboken: Wiley, 2010.

[2] 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.

[3] "LYSO Scintillation Crystal", *LUXIUM solutions*, 2022. [Online]. Available: https://www.crystals.saint-gobain.com/radiation-detection-scintillators/crystal-scintillators/lyso-scintillation-crystals. [Accessed: 2022-12-11].

[4] C. Park, A. Lee, R. Kim and J. Moon, "Evaluation of the Detection Efficiency of LYSO Scintillator in the Fiber-Optic Radiation Sensor", *Science and Technology of Nuclear Installations*, vol. 2014, pp. 1-6, 2014.

[5] J. Billingsley, S. O'Keeffe, C. Fitzpatrick, E. Lewis and A. Al-Shamma'a, "A review of optical fibre radiation dosimeters", *Sensor Review*, vol. 28, no. 2, pp. 136-142, 2008.

[6] H. Zubair, M. Begum, F. Moradi, A. Rahman, G. Mahdiraji, A. Oresegun, G. Louay, N. Omar, M. Khandaker, F. Adikan, N. Noor, K. Almugren, H. Abdul-Rashid and D. Bradley, "Recent Advances in Silica Glass Optical Fiber for Dosimetry Applications", *IEEE Photonics Journal*, vol. 12, no. 3, pp. 1-25, 2020.

[7] "0.50 NA Step-Index Multimode Fibers", *Thorlabs*, c1999-2022. [Online]. Available: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=362&pn=FP1500URT. [Accessed: 2022-12-11].

[8] "Single Photon Detectors", *Thorlabs*, 2022. [Online]. Available: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=5255. [Accessed: 2022-05-04].

[9] M. Jelínek, B. Mikel and P. Zemánek, "Optical fibers forming to ionizing radiation sensors preparation", in *21st Czech-Polish-Slovak Optical Conference on Wave and Quantum Aspects of Contemporary Optics*, 2018, pp. 14-.

Industrial data logger

1st Ondřej Matula
Department of Control and Instrumentation, FEEC
Brno University of Technology
Brno, Czech Republic
xmatul31@vutbr.cz

2nd Tomáš Benešl
Department of Control and Instrumentation, FEEC
Brno University of Technology
Brno, Czech Republic
tomas.benesl@vutbr.cz

***Abstract*—This paper deals with the design of an ON-LINE diagnostic system named as "industrial data concentrator". The paper looks at the design itself in a comprehensive way and brings together parts such as defining useful process variables, and the actual design of the industrial data concentrator both in terms of hardware and software.**

***Index Terms*—Industrial data logger, ON-LINE diagnostic system, Technical condition monitoring, ESP32-S3, IEPE**

I. Introduction

Collecting and analysing data is a normal part of our everyday lives, but along with technological advances, data analysis has been massively applied in recent years in fields where it was not possible before. One of these fields is industrial automation, in which data collection and analysis now has an irreplaceable place. The tool that deals with this task is technical diagnostics. It deals mainly with the determination of the technical condition of equipment and machines, the choice of the determining diagnostic variable, the methodology of data measurement, the identification of faults and the optimization of the production process. [1]

However, the main aspect of why technical diagnostics is becoming increasingly popular in industry is primarily economic. Thanks to the outputs that technical diagnostics provides, it is possible to plan maintenance or replacement of machine parts at the optimal time and thus avoid financial losses, which is also the main motivation for the design of the technical diagnostic system called "industrial data concentrator" that is the focus of this paper. [1]

Such devices are already available on the market, but most of them (found in my research) are not economically acceptable enough to be deployed on a massive scale in industry. Therefore, the aim was to design an industrial data concentrator as an ON-LINE diagnostic system that continuously monitors and evaluates the technical condition of the equipment and will measure all relevant process variables to determine the technical condition of rotating machines (e.g. inducion motors). At the same time, an effort was made to select system parameters and component pricing to make the data concentrator competitive.

In practice, the industrial data concentrator plays the role of DAQs (Data Acquisition System). It is a device placed next to the diagnostic object to which external sensors measuring the relevant diagnostic variables are connected. The measured data is then processed inside the data logger and sent via the local network to a local database server.

II. Requirements

Before designing the actual circuit solution, the functional, hardware and software requirements for the device were determined based on a comparison of commercially available products (e.g. IFM VSE-series, Flir SV87-kit, Siemens simotics connect 400, Bently Nevada 2300-series, National Instrument cRIO/cDAQ, Dynapar OnSite) and practical tests of measuring process variables. The individual requirements for the data concentrator are listed below.

- Functional requirements:
  - sending the measured data to the database server,
  - of data processing (e.g. FFT (Fast Fourier Transform), STFT (Short-Time Fourier Transform), correlation analysis, ordinal analysis),
  - indication of the operating state of the electric motor (on/off),
  - for counting operating hours,
  - counting the start of electric drives,
  - non-invasive current measurement,
  - multichannel temperature measurement,
  - vibration measurement,
  - magnetic field measurement,
  - speed measurement.

From a hardware point of view, a data concentrator is a device designed for installation in a small switchboard located in the immediate vicinity of the technical object of interest. The device itself does not contain any integrated sensors, all converters of process variables are external. The interface of the data concentrator was designed to support universal external sensors IEPE (Integrated Electronics Piezo Electric), RTD (Resistance Temperature Detector), digital in, digital out) but for integration reasons some interfaces were chosen as proprietary (e.g. interface for the current transformer). More detailed hardware and software requirements are listed below.

- Hardware requirements:
  - 24 VDC power supply,
  - optional 230 VDC power supply,

DOI 10.13164/eeict.2023.93

- ESP32-S3-WROOM-1 microcontroller,
  - Ethernet communication interface,
  - RS-485/PROFIBUS expansion interface,
  - IEPE interface for piezoelectric accelerometers (one channel),
  - temperature sensor interface RTD, NTC, PTC, KTY (two channels),
  - CT (Current Transformer) current measurement interface (three channels),
  - power digital outputs (two channels),
  - isolated digital inputs/outputs (four channels each),
  - analogue inputs (three channels),
  - EMC (Electromagnetic Compatibility) compliant,
  - DIN (Deutsches Institut für Normung) rail mounting.
- Software requirements:
  - using FreeRTOS (Real Time Operating System),
  - configuration via internal web server,
  - sending measured and processed data to local MySQL server.

As defined above, the industrial data concentrator should be connected to the local network and be able to send processed data to the database server. The way this has been designed is shown in Figure 1.

III. CIRCUIT SOLUTION

The circuit design is the main and most important part of the submitted work, based on the defined requirements.The device is divided into functional blocks shown in the block diagram in Figure 2. The individual parts are discussed below.

A. Microcontroller

The microcontroller was chosen with regard to the minimum price in relation to the required parameters, availability (as well as other components discussed in the paper) but above all due to the author's previous experience with this microcontroller. With this in mind, the microcontroller was selected from the Espressif catalogue, specifically the ESP32-S3-WROOM-1U-N16R2. It is a dual-core 32-bit microcomputer whose cores can operate at a frequency of up to 240 MHz. The multi-core solution is important mainly in terms of data processing, where one core can process the measurements and the other can send the data to the server. Another important feature is the support for extended instructions dedicated to FFT that support the butterfly algorithm. The microcontroller also has internal hardware support for IEEE 802.11 b/g/n Bluetooth LE v5.0 wireless protocols, which are used in the initial configuration.

Individual peripherals can be mapped to almost any physical GPIO (General Purpose Input Output) microcontroller using a GPIO matrix (slow speed), or an IO MUX (direct fast connection). This allows for great variability in the choice of peripherals.

The suffix ”U” in the chosen microcontroller model series indicates the version compatible with IPEX antenna. The antenna is integrated on the integrated PCB (Printed Circuit Board) as standard, it was chosen with regard to parasitic interference in industrial environment and therefore the need to use metallic (shielded) mechanical construction. [3]

B. Power supply

The input supply voltage has a choice of industry standard 24 VDC or direct line voltage from 85 VAC to 305 VAC. The power circuit topology is multi-stage due to the requirement for a large variety of voltages for peripherals.

The power distribution is mainly made up of switching power supplies due to power dissipation, but for analog power circuits the power supply branches are supplemented with appropriate filters and linear regulators so that the output of the switching power supplies is not subject to interference. The galvanic isolation of the power supply circuit must be located in the external 24 VDC source, if the power is supplied via the optional 230 VAC input the galvanic isolation is internally provided. The power circuits shall also include a 24 VDC isolated power supply to meet the needs of galvanically isolated digital inputs and outputs.

Of particular importance for the power supply is the correct connection of the grounds of the individual power circuits, therefore the device has been implemented on a four-layer PCB where the power and ground can be routed in separate layers.

C. Current measurement

The interface for current transformers is universal and depends only on the choice of the burden resistor in the specific circuit.

First, the signal from the current transformer is converted to voltage by the burden resistor, which is chosen with respect to the maximum output voltage. Next, the signal is fed to a precision rectifier to double the dynamic range of the measured current. [2] Operational amplifiers operate with asymmetrical supplies of 0 VDC to 3.3 VDC, so it was necessary to select a rail-to-rail amplifier (MCP6244). The precision rectifier has a gain of 0 dB therefore a non-inverting amplifier with selectable gain was included behind it. The entire current transformer interface cascade is terminated by a second order anti-aliasing filter (-40 dB/dec) of Sallen-Key topology.

An internal ADC of the microcontroller is used to digitize the analogue quantity from the transformer, which has a resolution of 12-bit and a sampling rate (RTC controller) of 100 kS/s before the multiplexer. The wiring is chosen so that one ADC (Analog To Digital Converter) is in charge of three current transformer inputs. The set sampling frequency is therefore 10 kHz, which allows, according to Nyquist's theorem, to measure a frequency up to 5 kHz, so that no aliasing effect occurs. The internal ADC has a programmable attenuator option which allows the measurement range to be changed programmatically.

D. Temperature measurement

Temperature measurement is one of the key process variables in industrial data application concentrator, so it must meet the criteria for accuracy and reliability of measurement.

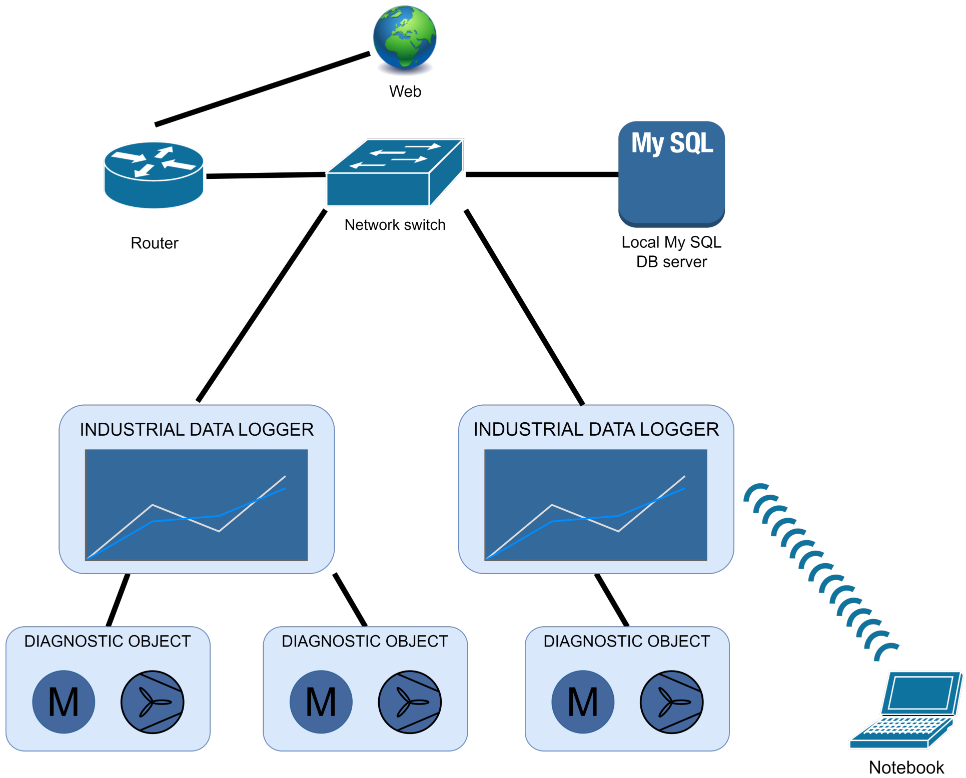


Fig. 1. Demonstration of integration into the local network system.

In the search for an efficient solution, the resistive principle was finally chosen for integration, combining good accuracy, temperature range and affordability. Finally, a dual-channel variant was proposed for temperature measurement. Analog Devices MAX6657MSA+ circuitry was used. The circuits integrate 15-bit ADC and associated input circuitry, so that optional support for two, three and four wire resistive sensor connections. The circuits declare the accuracy of 0.5 °C and a resolution of 0.125 °C. Various nominal sensor resistance values (PT100, PT1000, NTY, NTC) but the circuit must be reconfigured programmatically. [4]

E. Vibration measurement

Vibration measurement is the most important process variable that the data concentrator allows to measure. Measurement is possible either via the analogue IEPE interface (accelerometers based on the piezoelectric principle) or via the PROFIBUS to which a MEMS (Micro Electro Mechanical Systems) accelerometer can be connected.

The IEPE interface works on the principle of changing the internal resistance of the sensor depending on the magnitude of the deflection, this is accompanied by a change in the output voltage from the sensor.

The interface therefore consists of a programmable current source in the range of 2 to 20 mA with a voltage limitation up to 30 VDC. This is provided by the XTR111 circuit which operates in CC (Constant Current) mode depending on the input voltage from the DAC5311 circuit. The input is complemented by a control of the disconnected and shorted circuit using two comparators.

The signal is then stripped of the DC component by a coupling capacitor and fed into an attenuator consisting of a resistor network which provides attenuation of -6 dB, -12 dB, -18 dB and -24 dB. It is not possible to programmatically change the attenuator transmission, so a non-inverting amplifier with programmable gain of 6 dB, 12 dB, 15 dB and 18 dB was added (changing the value of the feedback resistor using the TS3A5017DR programmable analogue multiplexer).

The signal then passes to a differential anti-aliasing filter, which provides greater immunity to consensual interference than standard circuitry. The wiring is of the Sallen-Key type and is supplemented by a standard passive low-pass filter. The filter is third order, and the pass band is set to 20 kHz.

The filter output is further connected to a 24-bit sigma-delta ADC AD7768 with a sampling rate up to 1024 kS/s. The ADC has been chosen with regard to component availability, so the sampling rate will be limited to approximately 400 kS/s so that the input signal is sampled at 20 samples per period. This limited sampling rate, as opposed to the maximum, also gives me the ability to implement two measurement channels by multiplexing into one ADC in a future version of the data concentrator. With respect to the required data flow from the ADC, the respective buses were dimensioned.

Setup and data readout from the ADC is done programmatically via the SPI (Serial Peripheral Interface). The converter also includes a REF6025 voltage reference with a nominal

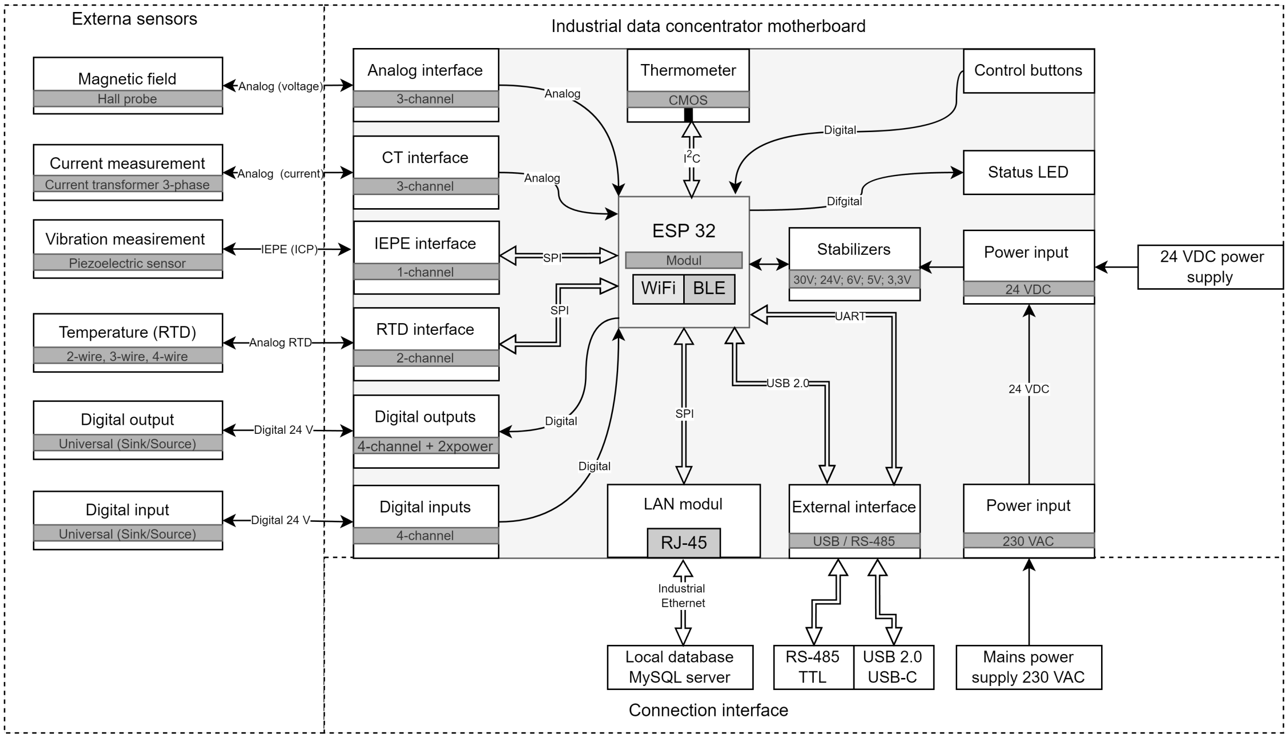


Fig. 2. Block diagram of industrial data concentrator.

output voltage of 2.5 V and a low temperature drift of 5 ppm/°C. The converter is clocked by an external 16.384 MHz oscillator. [5]

F. Communication interface

The communication interface mediates communication between the MCU (Micro Controller Unit) and external devices. Within the data concentrator, the Ethernet is used as the primary communication interface and Wiznet W5500 circuit is used as the transmitter. The interface is isolated using coupling transformers integrated into the RJ-45 input connector.

As mentioned in the introduction, the device has a Wi-Fi interface, but for stability reasons it is intended for initial configuration. In case the device cannot be configured wirelessly, a USB 2.0 configuration interface with USB-C connector is added. However, such configuration is limited and primarily intended for defining the network parameters of the device.

The external RS-485/PROFIBUS expansion interface is isolated (using iCoupler technology) and communicates in half-duplex mode. The communication is handled by the ADM2486BR transmitter. This is complemented by a terminator switch and also an isolated 5 V output power supply for the other transmitters on the link. The interface can communicate at up to 20 Mbps.

IV. Conclusion

In the paper I dealt with the complex design of industrial data concentrator as an ON-LINE diagnostic system. The comprehensive design consisted of a survey of current commercially available solutions and then drawing implications in parameterizing the requirements for the data concentrator. At the same time, relevant process variables were selected with respect to the technical diagnostics of rotating machinery.

The main part of the work was to design a circuit solution to meet the defined requirements and this was implemented. The circuit diagram and the printed circuit board were designed. Now the device is in the production phase and testing will be carried out.

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] HELEBRANT, František a Jiří ZIEGLER. *TECHNICKÁ DIAGNOSTIKA A SPOLEHLIVOST II: Vibrodiagnostika*. Ostrava, 2020.

[2] HOROWITZ, Paul a Winfield HILL. *The art of electronics*. Third edition. New York: Cambridge University Press, 2015. ISBN 978-0-521-80926-9.

[3] ESP32-S3. *ESP Prduct Selector* [online]. Maharashtra 411045,: Espressif Systems, c2015-2023 [cit. 2023-01-04]. Dostupné z: https://products.espressif.com/#/product-selector?names=

[4] MAX31865: RTD-to-Digital Converter. *Analog* [online]. Wilmington: Maxim Integrated Products, 2015 [cit. 2023-03-12]. Dostupné z: https://www.analog.com/media/en/technical-documentation/data-sheets/max31865.pdf

[5] AD7768-1: Analog devices. *Analog* [online]. Wilmington: Analog Devices, c2018—2019 [cit. 2023-03-12]. Dostupné z: https://www.analog.com/media/en/technical-documentation/data-sheets/ad7768-1.pdf

Speech generator for advanced embedded systems

Eliška Homzová
Department of Control and Instrumentation
FEEC Brno University of Technology
Brno, Czech Republic
xhomzo01@vut.cz

Petr Petyovský
Department of Control and Instrumentation
FEEC Brno University of Technology
Brno, Czech Republic
petyovsky@vut.cz

***Abstract*—This paper deals with a design and implementation of a speech generator for an advanced embedded system which converts written English text into synthetic speech. The first part describes the Atari 520ST Speech Synthesizer V2.0 program, which was the basis for the implemented speech generator. Then the design and implementation of the speech generator is described.**

***Index Terms*—embedded system, text-to-speech generator, voice generator**

I. Introduction

Speech generators are programs or devices that are designed to generate synthetic speech that is understandable and sounds natural [1]. Some speech generators produce synthetic speech so believable that they can mimic the voice and speech patterns of a real person [2].

This paper focuses on the generation of synthetic speech from text which we encounter almost everywhere today without even realizing it: from computer games to mobile phones or medical devices [1]. This type of generators is called a text-to-speech generators and consists of two main parts. The first part is natural language processing which analyses the input text and converts it into a phonetic transcription. In the second part, which is called speech synthesis, prosody and then synthetic speech are generated from the phonetic transcription [3].

This paper describes the implementation of a text-to-speech generator for an advanced embedded system. The basis for the implemented speech generator was the Atari 520ST Speech Synthesizer V2.0 program, which was created for the Atari 520ST personal computer. The Raspberry Pi Pico microcontroller board was chosen as an advanced embedded system for the realisation of the speech generator.

II. Atari 520ST Speech Synthesizer V2.0 program

Atari 520ST Speech Synthesizer V2.0 is a public domain program from 1986 by A. D. Beveridge and M. N. Day. The program was developed for the Atari 520ST personal computer and converts English text or phonetic transcription into synthetic speech [4].

A. Atari 520ST

The Atari 520ST is the first 16-bit personal computer in the Atari ST line of personal computers from the Atari Corporation and was released in 1985 [5].

The Atari 520ST contains the MC68000 microprocessor, the first implementation of the M68000 16/32-bit microprocessor architecture. It is also equipped with a YAMAHA YM2149 sound chip, formerly manufactured by General Instrument as a AY-3-8910. The YM2149 has three independently programmable tone generators which produce square waves, and a programmable noise generator. The outputs of the tone and noise generators are combined in a mixer. The YM2149 also has a programmable envelope and 15 exponential volume levels [5].

Also important to the program is the MC68901 multi-function peripheral chip from Motorola Semiconductor, which contains timers for program interrupts [4] [5].

B. Speech Synthesizer program description

As mentioned above, the Atari 520ST Speech Synthesizer V2.0 converts English text or phonetic transcription into synthetic speech. In addition to setting the type of input data, it is possible to set the tempo and intonation of the generated synthetic speech. The user can control the program using the program console [4].

Unfortunately, the original source texts of the public domain program are no longer available. Therefore, we were forced to perform a backward analysis of the program's machine code using the Ghidra tool. The original program .TOS file has a size of 28 kB and the converted analyzed code has a size of 195 kB in .gzf format. The analyzed code is written in the Assembly 68000 programming language.

The analyzed program can be divided into five main parts, as shown in Fig. 1.

After the program is started, the user can choose between generating synthetic speech from an English text and a phonetic transcription by typing the character "`.`" into the program console [4].

The user can also set the tempo of the synthetic speech by typing the character "`%`" and a value in the range from 20 to 199. When it is from the interval, the new tempo will be set. Otherwise the program reports an error [4].

A value of 20 corresponds to the fastest speech, a value of 199 corresponds to the slowest speech. The default value is 80 [4].

The intonation of the synthetic speech can be adjusted by entering the character "`!`" and a value in the range from 20 to 199 [4].

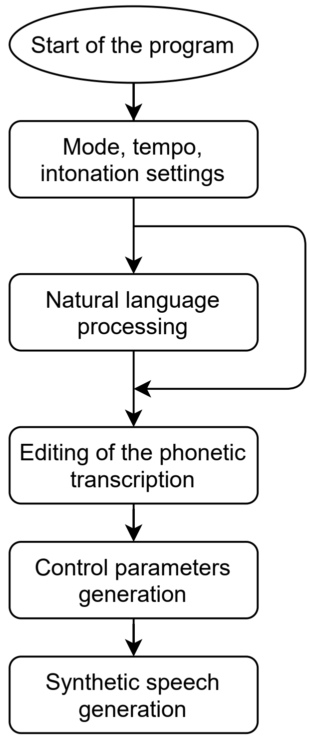


Fig. 1. A block diagram of the Atari 520ST Speech Synthesizer V2.0 program.

A value of 48 corresponds to shrill, a value of 80 corresponds to normal and a value of 199 corresponds to bass voice. The default value is 78 [4].

Synthetic speech generation starts by typing English text or phonetic transcription into the program console [4].

When synthetic speech generation from an English text is selected, natural language processing follows. Natural language processing begins by converting lowercase letters to uppercase letters. The actual conversion of the English text to phonetic transcription is done using a dictionary, which stores phonetic transcriptions of commonly used words, special characters and numbers from 0 to 9 [4].

All other words are converted using phonetic transcription rules, which are created by storing parts of words in a phonetic conditional table. The phonetic conditional table contains the parts of words, their phonetic transcription, and information about which characters can appear before and after a given part of a word to use a given phonetic transcription [4].

Natural language processing is followed by modifications to phonetic transcription, such as changing the tone and duration of the consonant according to the vowel that follows it [4].

Furthermore, additional phonemes are added after the phonemes of the diphthongs, or the phonemes are changed according to what comes before them. For example, the phoneme S is replaced by the phoneme Z if the phoneme G is in front of it [4].

Synthetic speech is generated sequentially by generating of the voiced and unvoiced sounds of each phoneme [4].

For sounds 6 control parameters are calculated in the Control parameters generation block from which the placement in the volume table is then calculated [4].

For unvoiced sounds the noise period for the YM2149 noise generator and the volume of the C channel to which the noise generator output is connected are also calculated [4].

The values for calculating of the control parameters are stored in a table with 16 values for each phoneme. Each phoneme has values indicating what type of phoneme it is, how many sounds it consists of, the period value of the noise generator, and values for calculating of the placement in the volume table [4].

The volume table stores 256 combinations of channel A, B and C volume levels, which when added together result in 256 volume levels. For each sound, its volume combination is found and written to the sound chip's registers. The program only uses the registers to set the channels volumes and the period of the noise generator. The tone generators periods are set to 0 [4] [6].

In this way, when the volumes of channels A, B and C are changed rapidly, their output is modulated and an 8-bit pulse code modulation is generated [6]. Fig. 2 shows a graphical overview of the resulting 256 volume levels converted to output voltage.

The frequency at which the volume registers of the sound chip are updated is determined by the set intonation value. The frequency at which the sounds are changed is determined by multiplying the tempo value by 77 and then dividing it by the intonation value [4]. The generated synthetic speech does not sound natural. However, it is mostly understandable.

III. Evaluation board with Raspberry Pi Pico

Raspberry Pi Pico is a microcontroller board based on the Raspberry Pi RP2040 microcontroller chip which contains a dual-core Arm Cortex-M0+ processor with a flexible system clock up to 133 MHz. It also includes 26 multi-function 3.3 V general purpose I/Os, 3 of which can be used as analog I/Os [7].

The Raspberry Pi RP2040 supports UART serial communication, SPI communication, I2C communication and has 16 PWM channels. It also has 2 MB of Flash memory and 264 kB of internal SRAM in six independent banks [7].

Raspberry Pi Pico was chosen because it is an affordable microcontroller board that supports the C/C++ programming language. And it can be extended with various circuits such as speakers.

Raspberry Pi Pico is connected to the evaluation board from the work on the link [8].

IV. Reimplementation of text-to-speech generator

The reimplemented speech generator is written in the C/C++ programming language and will form a software library for the selected embedded system.

The speech generator consists of 6 blocks and two branches, which are connected in the Synthetic speech generation block.

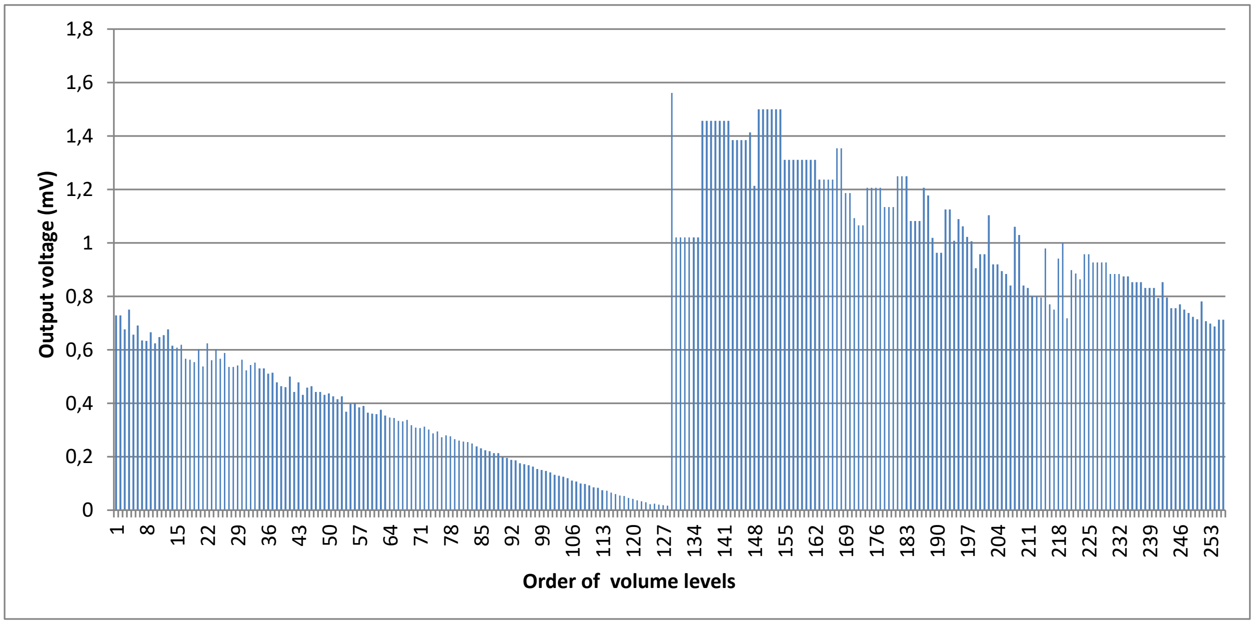


Fig. 2. A graphical overview of the resulting 256 volume levels converted to output voltage.

The first branch is the Tempo and intonation setting block, where new tempo and intonation values are verified and set.

Speech tempo and intonation are set independently, as the user may not want to change their values each time speech is generated. The tempo and intonation values can be set in the range from 20 to 199.

The second branch of the block diagram is the main algorithm of the speech generator, whose input values are the English text and the mode in which the main algorithm is to run. The block diagram of the reimplemented speech generator is shown in Fig. 3.

The reimplemented speech generator has 4 modes. In the first and second modes, synthetic speech is generated from the input data. In the first mode the input is an English text and in the second mode the input is a phonetic transcription. In the third mode, synthetic speech is generated from the previous input data, and the fourth mode is used to verify the completion of synthetic speech generation.

If the first mode is set, the Natural language processing block follows, in which the input English text is converted to phonetic transcription. The conversion to phonetic transcription is done in the same way as in the Atari 520ST Speech Synthesizer V2.0 program, using a combination of phonetic transcription rules and a dictionary that stores phonetic transcriptions of special characters, numbers and frequently used words.

The conversion of the input English text to phonetic transcription using phonetic transcription rules and a dictionary is done by going through a table in which the whole English words or their parts are stored together with their phonetic transcriptions. The table also contains information about which characters can appear before and after them.

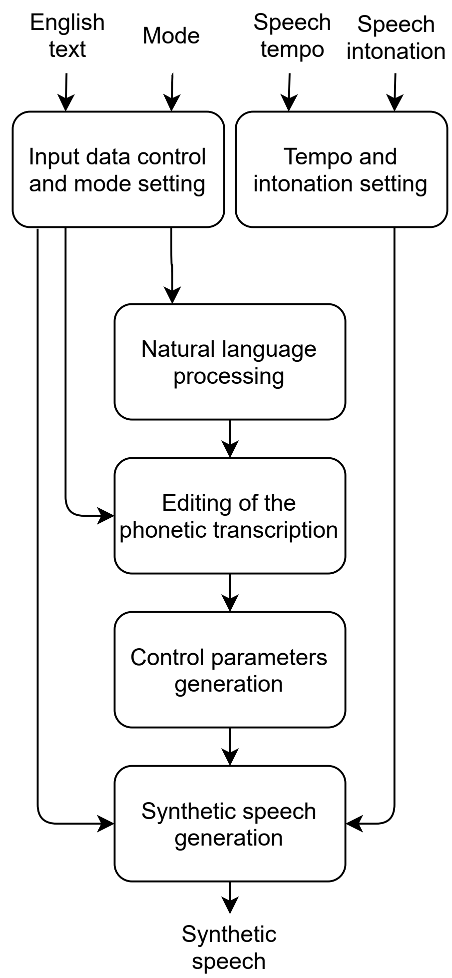


Fig. 3. A block diagram of the reimplemented speech generator.

For the first and second modes phonemic transcription modification follows, in which new phonemes are added to the phonetic transcription, or existing phonemes are modified, depending on where they are in the word.

The control parameters for the speech generator are calculated as in the Atari 520ST Speech Synthesizer V2.0 program from a table of values for calculating the control parameters for each phoneme. This table also contains information about how many sounds each phoneme consists of.

First, the values of the 6 control parameters are calculated. These parameters form the upper and lower bytes of the locations of 3 values in a volume offset table.

The values found in the volume offset table, when added together, form the location of the required channels volume levels combination in the volume table. The volume table contains the volume levels combinations for writing to channels A, B and C, as in the Atari 520ST Speech Synthesizer V2.0 program. Afterwards, the calculated values of the 6 control parameters are adjusted according to the following phoneme.

After calculating the 6 control parameters for each phoneme, the next step is to calculate the noise generator frequencies and its volumes for each sound. The noise generator is only used when generating unvoiced sounds, so for voiced sounds the value is 0. This makes possible to switch between voiced and unvoiced sound generation algorithms according to the set value of the noise generator frequency.

The noise generator consists of a 17-bit linear feedback counter, in which the feedback is the exclusive logical sum of bits 2 and 0. The output of the noise generator is the value of bit 1.

The generation of control parameters is followed by speech generation, which is performed by generating channels volume levels with a certain frequency. The frequency of sound generation is determined by the set speech intonation.

With this frequency, an interrupt is called to initiate the calculation of volume levels for the current sound. The calculated volume levels of the three channels are then added together, since the reimplemented generator has one output channel.

For unvoiced sounds, the volume of the noise generator is added instead of the volume level of channel C, but only when the output of the noise generator is equal to 1. Otherwise, nothing is added. The summed volume levels are then converted from the 0 to 45 range to the 8-bit PCM range.

The frequency at which the transition to the next sound is made is determined by the set speech rate, which is multiplied by 77 and then divided by the set speech intonation.

Currently, the functionality of the generator is tested by writing the output volume levels into .WAV format file.

V. Future work

In the following work, the speech generator will be modified and tested for use on a selected microcontroller board.

The timer of the Raspberry Pi Pico microcontroller will be used to generate the frequency of the noise generator and to determine the frequency of individual sounds generation.

The speech generator will have one output, which will use one of the PWM channels of the Raspberry Pi Pico microcontroller board.

As part of the improvements, the dictionary will be extended to include phonetic transcriptions of commonly used acronyms and numbers, so that the generator can pronounce numbers greater than 9. In addition, the volume table from the Atari 520ST Speech Synthesizer V2.0 will be modified, which may improve the intelligibility of the synthetic speech.

VI. Conclusion

The aim of this work was to design and implement a text-to-speech generator for an advanced embedded system. The basis for the our implemented speech generator was the Atari 520ST Speech Synthesizer V2.0 program, which provides synthetic speech generation from an English text and a phonetic transcription.

The Atari 520ST Speech Synthesizer V2.0 public domain program uses a dictionary and phonetic transcription rules to convert text to phonetic transcription. And it uses pulse code modulation to generate synthetic speech, a feature of the YM2149 sound chip not described in the datasheet.

The Raspberry Pi Pico microcontroller board was chosen as the advanced embedded system. The Raspberry Pi Pico was selected because it is an affordable microcontroller board that supports the C/C++ programming language.

The reimplemented speech generator has separate settings for intonation and tempo and can run in 4 modes. It also uses a dictionary and phonetic transcription rules to convert English text into phonetic transcription.

Once the program is complete and tested, it will be available as a software library.

References

[1] What is Speech Synthesis? A Detailed Guide. WebsiteVoice [online]. 24 August 2022 [cit.2022-12-10]. Available: https://websitevoice.com/blog/what-is-speech-synthesis/

[2] BROWN, Dalvin. AI gave Val Kilmer his voice back. But critics worry the technology could be misused.The Washington Post [online]. 18 August 2021 [cit.2022-12-10]. Available: https://www.washingtonpost.com/technology/2021/08/18/val-kilmer-ai-voice-cloning/

[3] HOLMES, John a Wendy HOLMES. Speech synthesis and recognition .2nd ed. London ; New York: Taylor & Francis, 2001, 298 p. ISBN 0-7484-0856-8.

[4] BEVERIDGE, A.D. a M.N. DAY. Atari 520ST Speech Synthesizer V2.0. Demozoo [online]. [cit.2022-12-16]. Available: https://demozoo.org/productions/125048/

[5] LENDINO, Jamie. Faster Than Light: The Atari ST and the 16-Bit Revolution. Audubon, NJ: Steel Gear Press, 2019, 325 p. ISBN 978-1732355217.

[6] SILVEIRA, Marcelo. 8-bit sounds on MSX PSG [online]. 2022, 11 p. [cit. 2022- 12-20].Available: http://marmsx.msxall.com/artigos/som_8bit_en.pdf

[7] RASPBERRY PI LTD. Raspberry Pi Pico Datasheet: An RP2040-based microcontroller board [online]. Version 2.1. 2023, 31 p. [cit. 2023-03-03]. Available: https://datasheets.raspberrypi.com/pico/pico-datasheet.pdf

[8] PONČÁK, Matej. Inovace laboratorních úloh kurzu Vestavné systémy [online]. Brno, 2022 [cit. 2023-03-10]. Available: http://hdl.handle.net/11012/204867. Diplomová práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav automatizace a měřicí techniky. Vedoucí práce Petr Petyovský.

Sparse Representation for Classification of Posture in Bed

1st Michaela Mesárošová
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
xmesar01@vut.cz

2nd Ondrej Mihálik
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
ondrej.mihalik1@vut.cz

***Abstract*—Redundant dictionaries, also known as frames, offer a non–orthogonal representation of signals, which leads to sparsity in their representative coefficients. As this approach provides many advantageous properties it has been used in various applications such as denoising, robust transmissions, segmentation, quantum theory and others. This paper investigates the possibility of using sparse representation in classification, comparing the achieved results to other commonly used classifiers. The different methods were evaluated in a real-world classification task in which the position of a lying patient has to be deduced based on the data provided by a pressure mattress of 30×11 sensors. The investigated method outperformed most of the commonly used classifiers with accuracy exceeding 92%, while being less demanding on design and implementation complexity.**

***Index Terms*—sparse representation, linear regression, LASSO, redundant basis, SRC, classification**

I. Introduction

Signals are often represented in different systems, typically bases, which provide a non-redundant signal representation. Any corruption or loss of transform coefficients can lead to serious information loss about the original signal. In environments in which the integrity of the signal representation cannot be reliably guaranteed, redundant systems provide useful properties to combat the threat of information loss. Furthermore, the popularity and applications of redundant systems have been increasing over the last years [1]–[3]. This paper focuses on the application of redundant dictionaries for signal processing. We verify and assess the possibility of using this representation in classification tasks, in which the position of a person lying on the pressure-sensitive mattress needs to be determined.

The outline of this paper is as follows. Section II introduces the necessary mathematical apparatus. Section III provides an analysis of the dataset with its representation using redundant coefficients. Section IV discusses the classification criteria and their results. Lastly, in Section V, all achieved results are compared against commonly used classifiers.

II. Linear regression models

As stated in [4], linear regression has been widely used for decades and still constitutes an important pillar in statistics. It is simple and often provides easily interpretable description of the relation between the inputs and output of an analysed system. In general, regression uses a pool of mathematical and statistical tools to determine the estimation of the relationship between variables. Specifically, we can consider the general relationship among the independent variables $\mathbf{x}$, the dependent variables y, and the vector of unknown parameters β:

$$y = f(\mathbf{x}, \beta) \tag{1}$$

where f represents regression function, typically prescribed, and β are parameters found through optimization of the goodness-of-fit of this function. [5] Successful optimization is immensely dependent on the expression we want to optimize. As an input we assume p discrete linear inputs

$$\mathbf{x} = \{x_1, x_2, \ldots, x_p\}^T \tag{2}$$

and we aim at the prediction of the output y. Thus, the regression model has the following form:

$$\hat{y} = \hat{\beta}_0 + \sum_{j=1}^{p} x_j \hat{\beta}_j \tag{3}$$

where $\hat{y}$ is marking predicted value and $\hat{\beta}_0$ represents the bias [4].

A. LASSO

Generally, regression model is more interpretable, if it has fewer descriptive coefficients that bear on the outcome. Least absolute shrinkage and selection operator (LASSO) adds an ℓ_1 penalty term to regularize the least-squares regression. As a result we are able to effectively penalize outliers and alleviate overfitting of the model to given data. Thus we get the following formula:

$$\hat{\beta} = \arg\min_{\beta} \|\mathbf{X}\beta - \mathbf{y}\|_2^2 + \lambda\|\beta\|_1 \tag{4}$$

where λ parameter directly affects sparsity of a resulting model. By increasing λ we restrict the outcome to fewer coefficients. However, decreasing this parameter allows more terms in the estimated model until the obtained result corresponds to the least-squares model when $\lambda = 0$ [5], [6]. In order to determine the optimal value of λ, we conducted a numerical experiment as follows. Squared sum of the residuals of the outcome and the time needed for computation of the model was compared for different λ values. Results are plotted in Fig.

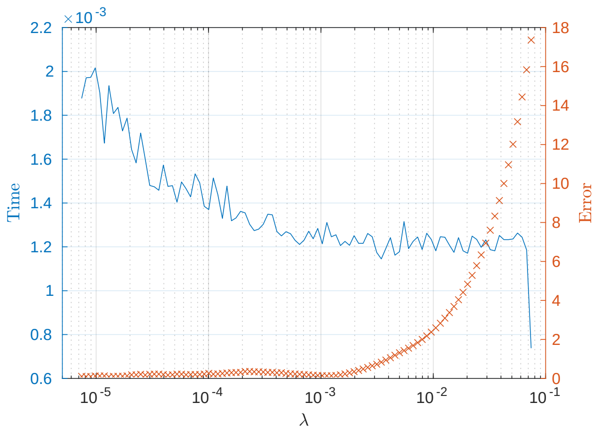


Fig. 1. Dependency of computational complexity and model error on λ

1. Based on the results, we chose $\lambda = 5 \times 10^{-3}$, for the model error lies in the stable interval, and there is no appreciable change in computational time when λ is increased further. In other words, the advantageous properties of a sparse model are retained, yet we do not have to sacrifice the estimation precision of the model. The same regularisation constant value is being used for the remainder of the paper.

III. Sparse representation of measurements

Signals analyzed in this paper are measurements obtained from a pressure-sensitive mattress stored as 30×11 matrices. Information associated with each observation is subject number and position number. Subject labels measured data by participating person whilst position number distinguishes between four possible lying positions, namely:

- on the back – class 1,
- on the left side – class 2,
- on the right side – class 3,
- on the stomach – class 4.

The dataset is similar to the set presented in [7], yet larger. It comprises 18 different test subjects in varying positions, containing 290 pressure map images in total. For the sake of clarity, we present a few data examples in Fig. 2.

The strategy chosen to train classification models and evaluate their accuracy is cross-validation [4]. Data are split into folds; each fold comprising pressure images of one individual. This yields 18 folds: 17 are used for training and one for validation. Therefore, there are 18 possible ways of data partition, which yield 18 classifiers of the same type, each validated using a different fold.

Training images are reshaped into column vectors, normalized to unit ℓ_2 norm, and stacked into a new matrix so as to form the redundant dictionary $\mathbf{X}$. The dictionary is appended by an identity matrix, which affords and effective means of dealing with occlusions [3], [7].

The first part of the algorithm consist in decomposing a validation signal by means of (4), thereby obtaining sparse coefficients $\hat{\beta}$, which form the signal representation as seen in Fig. 3. Although the signal decomposition may consist of nearly 250 observations, only a few were recognized by the LASSO algorithm as relevant to description and recreation of the validation signal. Furthermore, the dominant coefficients (blue) suggest that the image belongs to the first class. However, the the biggest coefficient (yellow) may signify that the observation belongs to the third class.

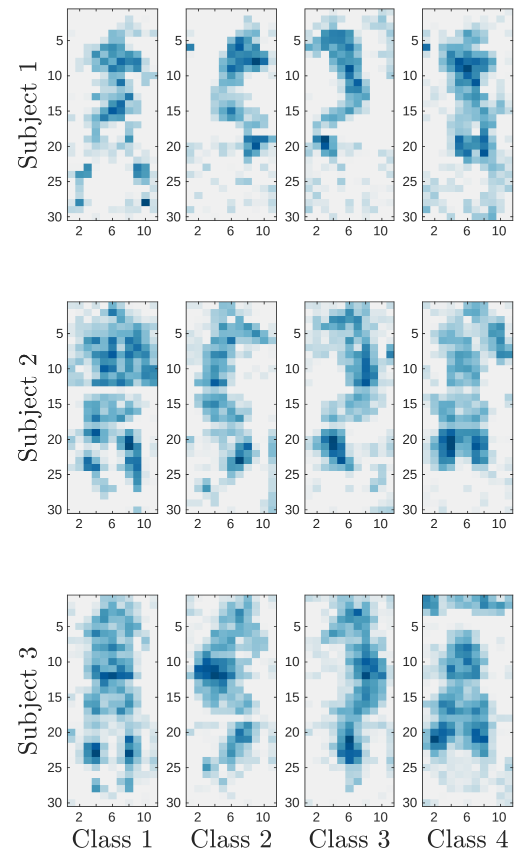


Fig. 2. Data visualisation for three exemplar subjects in all four possible positions

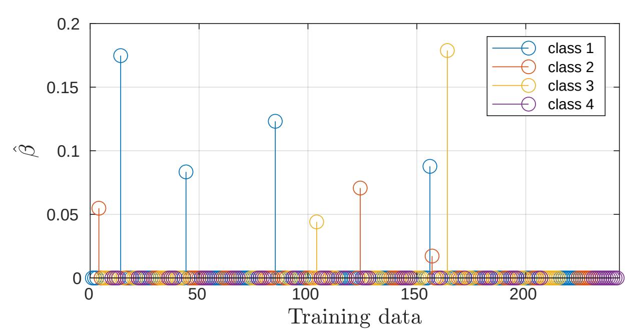


Fig. 3. Sparse coefficients for measurement 20 belonging to subject No. 1 from class No. 1

IV. Sparse representation classification

To be able to tell which measurement corresponds to which class, marked as k, it is crucial to find suitable decision-making scheme. To address this problem we adopted three methods presented in paper [8]:

- Maximum coefficient (MC) – The predicted class label corresponds to the class label of the training sample with the largest coefficient of the sparse solution

$$\hat{k} = \arg\max_{k}(\max_{i\in\mathbf{x}_{ki}} \beta_i). \tag{5}$$

- Maximum sum of class coefficients (MSCC) – The class whose sum of β coefficients is maximized is considered as the predicted class label. In other words, for each class k we take the sum of the coefficients β, and we classify the observation into the class with the largest sum.

$$\hat{k} = \arg\max_{k}\left(\sum_{i\in\mathbf{x}_{ki}} \beta_i\right). \tag{6}$$

- Minimum class residual (MCR) – This method predicts the class label $\hat{k}$ based on the value of class residual. Whichever class minimizes the residual is chosen as prediction label.

$$\hat{k} = \arg\min_{k}\|\mathbf{y} - \mathbf{X}_k\beta_k\|_2. \tag{7}$$

where $\mathbf{X}_k$ and β_k contain only observations and coefficients, respectively, corresponding to class k.

Calculation of all model coefficients with the cross-validation setup introduced in III allows creating reconstructions of original images as follows:

$$\mathbf{y} = \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \cdots + \mathbf{x}_p\beta_p \tag{8}$$

For the sake of clarity, the behaviour of (5), (6), and (7) will be illustrated on measurement No. 20 whose sparse coefficients are depicted in Fig. 3.

A. Maximum Coefficients

MC is the simplest out of the three methods introduced above. Fig. 4 shows value of maximal coefficient for each class separately, where it is very easy to unveil the disadvantages of the MC approach. Fig.3, despite the prevailing occurrence of first class coefficients, the decision is made only on the basis of one element. Other coefficients are not involved in the classification process.

B. Maximum Sum of Class Coefficients

It is obvious from Fig. 4, that MSCC has successfully eliminated the unfavourable effect of a single strong coefficient, as in MC. The classification is correct and the difference between the MSCC values is larger, so the result is decided with more certainty.

C. Minimum Class Residual

MCR may be seen as a conceptual extension of the MSCC. If a class makes the highest contribution to forming the signal approximation, then a reconstruction image created by this class coefficients is likely to be the most similar to the original image. Therefore it can be expected for MCR and MSCC to deliver the same or very similar results of classification.

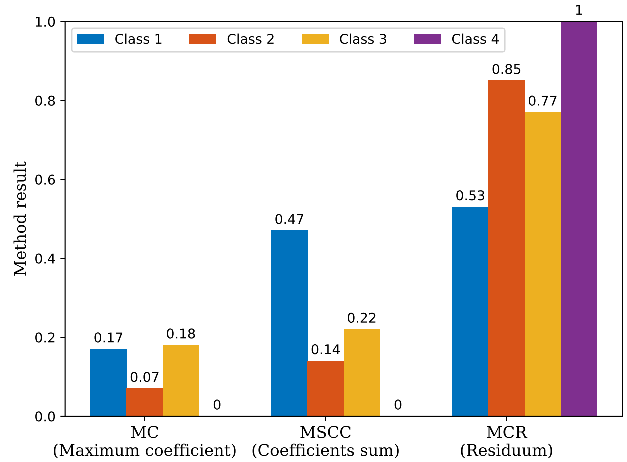


Fig. 4. Results obtained by MC, MSCC and MCR algorithms applied to each class separately using measurement No. 20 from subject No. 1 belonging to the 1st class.

The residual value reflects the error of the reconstructed observation; therefore, classes with zero-valued coefficients will reach 100 % reconstruction error as shown in Fig. 4.

D. Occlusions

In various applications there is a possibility of malfunctioning equipment or corrupted data. Therefore it is important to investigate behaviour of SRC in such cases. We will consider the fault being missing rows as it is a consequence of a non-functional sensor. To make classification possible despite errors in the data, new elements are appended to the dictionary $\mathbf{X}$ representing potential errors [7]. However the fault definition is interchangeable if a user has more information about typical fault character in specific case. Classification with modified dictionary was applied on original as well as on occluded data and results are listed in Table I in the subsequent section.

V. Result Comparison with Other Classifiers

To evaluate classification results for the whole dataset, algorithm is run in the manner described in Section III. Overall success rate is judged based on the TPR value.

$$\text{TPR} = \frac{\text{Correctly classified into the class}}{\text{Total number of class members}} \tag{9}$$

Global TPR value is conveniently calculated from the confusion matrices, see Fig. 5, where partial classification rates for each approach and class is displayed. The global TPRs obtained by MC, MSCC, and MCR are summarized in Table I. To evaluate the viability of SRC more objectively than in [8], the results obtained were compared to several commonly used machine learning approaches and classification methods: Convolutional neural network (CNN), K-nearest neighbours (KNN), linear discriminant analysis (LDA) and bagged classification trees, to name a few.

Most of the classifiers allow selection of a training parameter, such as the regularisation γ for (LDA). A broad range of

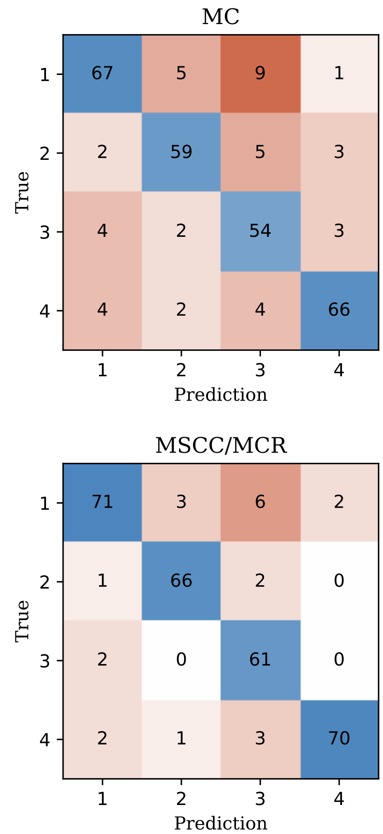


Fig. 5. Confusion matrix overview of the partial classification success. Since MSCC and MCR yield the same results we included them in one common graph.

parameters was tested and the classifiers yielding the highest validation TPRs were used in the table.

Quadratic discriminant analysis (QDA) could not be trained using the dataset without occlusions because the training algorithm failed due to singular covariance matrices. The CNN requires a proper selection of its structure. The following layers turned out to yield the best validation TPRs:

TABLE I
CLASSIFICATION SUCCESS RATES

Classification Approach	**TPR**			
	trained using clean dataset		trained to deal with occlusions	
	original validation data	occluded validation data	original validation data	occluded validation data
SRC, MC	0.8483	0.8303	0.8552	0.8502
SRC, MSCC	0.9241	0.9136	0.9172	0.9177
SRC, MCR	0.9241	0.9136	0.9172	0.9177
CNN (2 covolutional layers)	0.9310	0.9183	0.9344	0.9232
LDA, $\gamma = 0.76$	0.9034	0.8945	0.8483	0.8389
QDA, $\gamma = 0$	—	—	0.7897	0.7882
regularised logistic regression	0.8931	0.8817	0.8414	0.8300
SVM, box constraint = 0.9	0.8897	0.8776	0.8207	0.8122
KNN, $K = 1$	0.8586	0.8455	0.8483	0.8486
KNN, $K = 10$	0.8379	0.8216	0.8621	0.8554
50 bagged classification trees	0.8207	0.8144	0.8759	0.8720
Sole classification tree	0.5207	0.5077	0.5621	0.5606

1) 3 convolutional filters of size 3×3,
2) leaky relu layer,
3) max pooling layer of size 2×2 with stride 2,
4) 6 convolutional filters of size 3×3,
5) leaky relu layer,
6) max pooling layer of size 2×2 with stride 2,
7) fully connected layer.

The proper structure, however, was established after an extensive search through different structures. Hyper-parameters such as the numbers and sizes of layers, filters and pooling layers were tried before the aforementioned structure was found.

VI. Conclusion

This paper discussed possibilities of signal representation via redundant basis and its potential in solving classification tasks – classification of the patients' lying posture. SRC performance was compared to standard classification methods. As SRC requires but one parameter selection and no feature extraction is necessary, the implementation complexity of SRC is reduced. The accuracy achieved is not sensitive to regularisation parameter in a broad range around optimal value whereas the other compared methods required careful parameter selection using cross-validation, had we tested their accuracy on test data, their TPRs would very likely be smaller. If occlusions are taken into account, performance of SRC is changed only marginally, delivering higher TPR values than most of the common classifiers tested in this study. SRC TPRs were outperformed only by the CNN, but the extra effort expended in the design of CNN is scarcely worth the marginal increase in TPR.

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] S. Mallat, *A wavelet tour of signal processing: the sparse way*, 3rd ed. Boston: Amsterdam: Academic Press; Elsevier, 2009.
[2] J. Kovacevic, A. Chebira, *An introduction to frames*, Foundations and Trends® in Signal Processing, Vol. 2, No. 1, pp. 1–94, 2008.
[3] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
[4] T. Hastie, R. Tibshirani, J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, 2nd ed. New York: Springer, 2009.
[5] S. Brunton, J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, Cambridge: Cambridge University Press, 2019.
[6] R. Tibshirani, "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society. Series B (Methododological)*, Vol. 58, No. 1, pp. 267-288, 1996.
[7] O. Mihálik, T. Sýkora, M. Husák, and P. Fiedler, "In-bed posture classification based on sparse representation in redundant dictionaries," *IFAC-PapersOnLine*, Vol. 55, No. 4, pp. 374-379, 2022.
[8] J. J. Liu, "Sleep posture analysis using a dense pressure sensitive bedsheet," *Pervasive and Mobile Computing*, Vol. 10, pp. 34–50, 2014.

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

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

Automated Python-Controlled Thermal Analysis of Rotor of Line-Start Permanent Magnet Synchronous Machine

Jakub Súkeník
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
xsuken00@vutbr.cz

***Abstract*—This paper deals with the thermal analysis of the rotor of the line-start permanent magnet synchronous machine (LSPMSM). First, a concept of a fully controlled parametric geometry creation using the Python programming language is introduced. The geometry is created using the FreeCAD software. Afterwards, the geometry is used to perform a thermal analysis using the finite element method (FEM) in the Ansys Workbench Software. This analysis is also controlled by Python. A thermal model of the analyzed rotor was also created using lumped parameter thermal network (LPTN). Finally, the resulting temperatures from the FEM-based analysis are used to verify the functionality of the lumped-parameter thermal network.**

***Index Terms*—finite element method, heat transfer, line-start permanent magnet synchronous machine, lumped-parameter thermal network, automated parametric geometry creation using Python programming language**

I. Introduction

Electric motors for industrial and domestic applications account for the largest share of electricity consumption. Due to the rising electricity prices, material, and environmental awareness, there is a growing interest in high energy-efficient rotating machines [1].

To achieve improved parameters such as better efficiency and higher power density of electrical machines, thermal analysis in parallel with the electromagnetic analysis is necessary. The thermal analysis of the electrical machines can be performed using the lumped-parameter thermal network (LPTN) or the finite element method (FEM). The first method is analytical, and the second one is numerical [2], [3].

However, before the thermal and electromagnetic FEM analysis, the model of the electrical machine must be designed using modelling software such as Inventor, Ansys SpaceClaim, or FreeCAD.

FEM is suitable for modelling complex shapes, with the additional advantage of accurate heat transfer and heat loss distribution within solid materials, but is hardware demanding and time-consuming. FEM is usually used as a reference to LPTN analysis [4], [5].

The advantages of LPTN compared to FEM include fast calculation with acceptable accuracy. This method is suitable for quickly evaluating the temperature of the main components of electrical machines. For the LPTN, the two-dimensional heat conduction is usually considered [6].

The first part of the paper describes the heat transfer fundamentals. The second part presents an automatic creation of a parametric geometry using Python programming language. Furthermore, thermal rotor model of the LSPMSM using FEM is presented and the input data are introduced. Finally, results from the rotor thermal model of the LSPMSM using FEM and results from the rotor LPTN of the LSPMSM are compared.

II. Heat transfer

There are three mechanisms of heat transfer, namely conduction, convection and radiation. The temperature difference is balanced as heat is transferred from a warmer place to a colder place according to the second law of thermodynamics. These types of heat transfer occur simultaneously in many cases [7]. However, radiation heat transfer can be neglected in many cases and thus it is not considered in this paper.

The temperature difference between the outer surface of an electrical machine and the surroundings is usually about 30 K. Then, the heat transfer coefficient by radiation is approximately equal to 4 $\frac{\mathrm{W}}{\mathrm{m}^2\cdot\mathrm{K}}$. This value is very small compared to the heat transfer coefficient by forced convection [8].

A. Conduction heat transfer

According to [7], the fundamental relation for heat transfer by conduction in matter is Fourier's law, which is in the form of

$$\dot{\mathbf{q}} = -\lambda \nabla \vartheta, \tag{1}$$

where $\dot{\mathbf{q}}$ is the heat flux vector, λ is the thermal conductivity, $\nabla\vartheta = \left(\frac{\partial\vartheta}{\partial x}, \frac{\partial\vartheta}{\partial y}, \frac{\partial\vartheta}{\partial z}\right)$ is the temperature gradient. The negative

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 Mr. Marek Toman as my research supervisor for his valuable comments and help while writing this paper.

sign in (1) represents the orientation of heat flow in the direction of decreasing temperature [7].

The temperature field is obtained by solving the differential equation at the given boundary conditions. The general differential equation of non-stationary heat conduction in the Cartesian coordinate system has according to [7] the form

$$\frac{\partial}{\partial x}\left(\lambda_{\mathrm{x}}\frac{\partial \vartheta}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{\mathrm{y}}\frac{\partial \vartheta}{\partial y}\right)+\frac{\partial}{\partial z}\left(\lambda_{\mathrm{z}}\frac{\partial \vartheta}{\partial z}\right)+\cdots$$
$$\cdots+p_{\mathrm{gen}}=\rho c\frac{\partial \vartheta}{\partial t}, \quad (2)$$

where ϑ is the temperature, λ_{x}, λ_{y}, and λ_{z} are the thermal conductivities of the substance in the respective directions of the Cartesian coordinate system, p_{gen} is the internal loss generated in the substance per unit volume, ρ is the density, c is the specific heat capacity and t is the time. Equation (2) has to be solved numerically in most cases because of its difficulty to solve analytically.

B. Convection heat transfer

The amount of heat transfer between a solid body and surrounding fluid determines Newton's law of cooling according to [7] in the form

$$\dot{Q}=\alpha A\left(\vartheta_{\mathrm{s}}-\vartheta_{\infty}\right), \quad (3)$$

where $\dot{Q}$ indicates the heat transfer rate, A is surface of a body, ϑ_{s} is the temperature at the surface of the body, and ϑ_{∞} is the ambient temperature.

III. Creating automated parametric geometry using Python programming language

Creating geometry manually can be time-consuming and inconvenient when geometric dimensions change frequently. Therefore, a procedure was developed to automate the creation of geometry based on input parameters. Parametric geometry is important for FEM thermal analysis.

A parametric geometry is realized in FreeCAD. This software is controlled by Python code that FreeCAD runs as a subprocess. The Python code runs in the PyCharm environment. The detailed structure of the Python code for creating the model is discussed in section III-A.

The steps described below were performed with version 0.18.3 of FreeCAD and version 3.6.0 of Python [9], [10].

A. Parametric geometry creation algorithm

The algorithm for generation the parametric geometry is depicted in Fig. 1. First, the required libraries are loaded, and FreeCAD is started as a sub-process. Next, variables from another Python file are imported. This file contains the basic geometric dimensions and settings.

The following sequence of actions creates the parametric geometric model using a set of commands. Examples of these commands include a pocket, extrusion, line segment, and circle. The list of commands is situated directly in FreeCAD.

For more variability, the created parametric geometry needs to be exported. For this purpose, using the STEP (Standard for the Exchange of Product Data) format is performed. After this sequence, the program is finished.

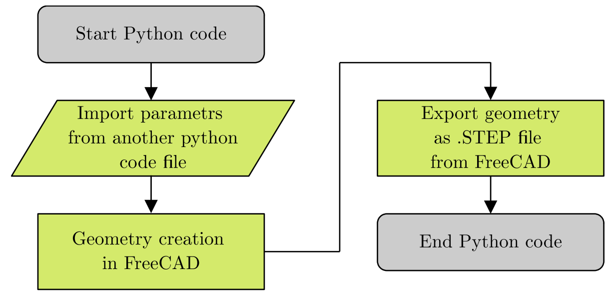


Fig. 1. Parametric geometry creation algorithm.

B. Parametric rotor geometries

In synchronous machines such as LSPMSMs, the magnet in the rotor can be demagnetized due to high temperatures. Therefore, it is important to simulate the thermal behavior of the rotor [11]. For this reason, the parametric rotor geometry for FEM analysis is necessary.

For more versatility of the rotor parametric geometry, the fourteen types of rotor slots were implemented in the developed script. The rotor slots are shown in Fig. 2 a) to n).

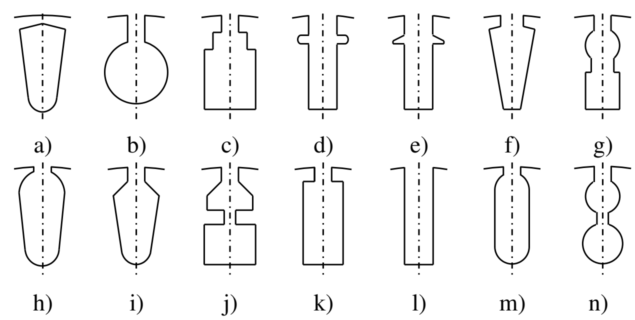


Fig. 2. Types of rotor slots.

The number of rotor slots can be arbitrarily chosen. All rotor slots can be adjusted to the required geometric dimensions.

In Fig. 3, the example of the proposed geometry of the LSPMSM parametric rotor for FEM analysis is depicted.

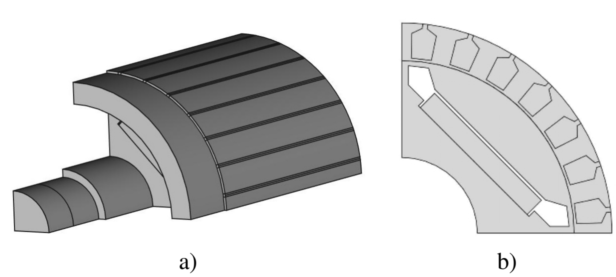


Fig. 3. Example of generated rotor geometry of LSPSMS.

IV. Thermal Analysis of Rotor of LSPMSM

The thermal analysis of the LSPMSM rotor is performed by FEM in Ansys Workbench software. The steady-state and transient thermal analysis is performed. Emphasis has been placed on automating the entire thermal analysis process.

A. Algorithm for the FEM thermal analysis of the rotor of the LSPMSM.

The algorithm for the FEM thermal analysis of the rotor of LSPMSM is described in Fig. 4. The algorithm consists of two branches. The right branch represents the creation of the thermal analysis and the left branch consists solution sequence.

In the thermal analysis block, the boundary conditions are automatically assigned to the proposed geometry.

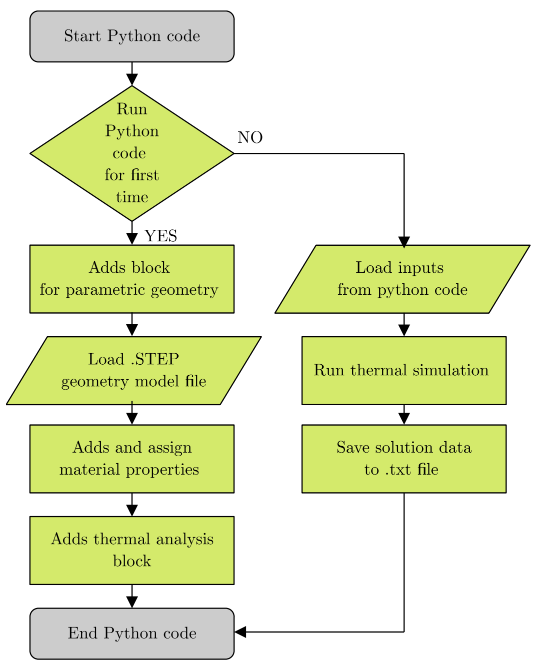


Fig. 4. Algorithm for the FEM thermal analysis of the rotor of LSPMSM.

B. Inputs entered into the LSPMSM rotor analysis

A thermal network of the LSPMSM rotor has been created and further compared with the FEM thermal analysis.

The rotor thermal network of the LSPMSM in Fig. 5 consists of eight nodes. At each node, the average temperature of a specific part of the LSPMSM rotor is calculated.

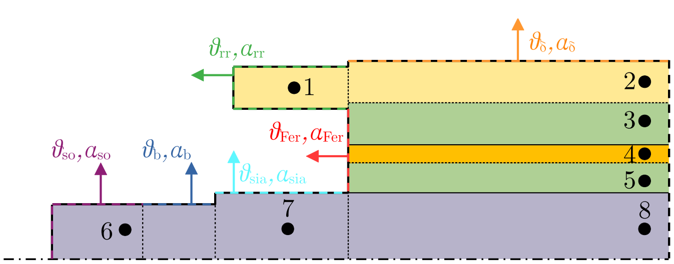


Fig. 5. Nodes distribution and convection boundary conditions.

In Fig. 5, convection boundary conditions are shown. Internal heat generation is not considered in the LSPMSM rotor thermal network.

In Table I, the convective heat transfer coefficients for the LSPMSM rotor thermal network are calculated from the peripheral rotor speed. These calculations are described in [8], [12]. Table II presents the material properties for the LSPMSM rotor thermal analyses.

TABLE I
The description of the nodes and convection heat transfer coefficients for the LSPMSM rotor thermal analyses

Node	Description	Convection heat transfer coefficient		
		Symbol	*Value*	*Unit*
1	Rotor end ring	α_{rr}	21	
2	Bars + teeth	α_{δ}	178	
3	Rotor upper yoke	α_{Fer}	21	
4	Permanent magnet	α_{Fer}	21	
5	Rotor lower yoke	α_{Fer}	21	$\frac{W}{m^2 \cdot K}$
6	Shaft (ambient air)	α_{so}	15	
		α_{b}	1415	
7	Shaft (inner air)	α_{sia}	27	
8	Shaft (inside rotor)			

TABLE II
Material properties for the LSPMSM rotor thermal analyses

Material[a]	Thermal conductivity		Density		Specific heat capacity	
	Value	*Unit*	*Value*	*Unit*	*Value*	*Unit*
Aluminium	240		2700		900	
Electrical steel[b]	35; 0.6	$\frac{W}{m \cdot K}$	7600	$\frac{kg}{m^3}$	450	$\frac{J}{Kg \cdot K}$
NdFeB	9		7600		425	
Steel	45		7700		500	

[a]The material properties come from [8].
[b]The first thermal conductivity is in the radial direction, while the second thermal conductivity is in the axial direction.

C. Steady-state analysis results

Fig. 6 shows the steady-state temperature field observed by FEM analysis performed in Ansys Workbench software.

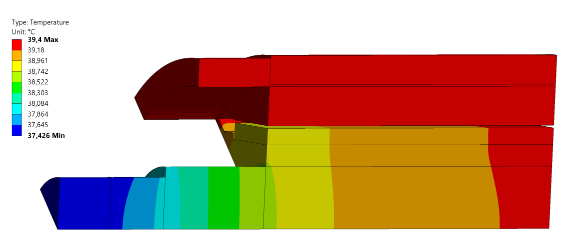


Fig. 6. FEM temperature field distribution.

Table III shows the average LSPMSM rotor steady-state temperatures using FEM analysis and the average temperatures using LPTN. The average temperature deviations in the form $\Delta\vartheta = |\vartheta_{FEM} - \vartheta_{LPTN}|$ between FEM and LPTN are very minor.

TABLE III
AVERAGE STEADY-STATE TEMPERATURES OF FEM AND LPTN AND THEIR COMPARISON

Node	FEM	LPTN	Difference
	ϑ_{FEM} (°C)	ϑ_{LPTN} (°C)	$\Delta\vartheta$ (K)
1	39.35	39.43	0.08
2	39.30	39.36	0.06
3	39.30	39.35	0.05
4	39.09	39.22	0.13
5	39.09	39.12	0.03
6	37.61	37.72	0.11
7	38.35	38.49	0.14
8	39.06	38.93	0.13

D. Transient analysis results

Transient rotor temperatures of the LSPMSM using the LPTN and average temperatures using the transient FEM in each node are shown in Fig. 7.

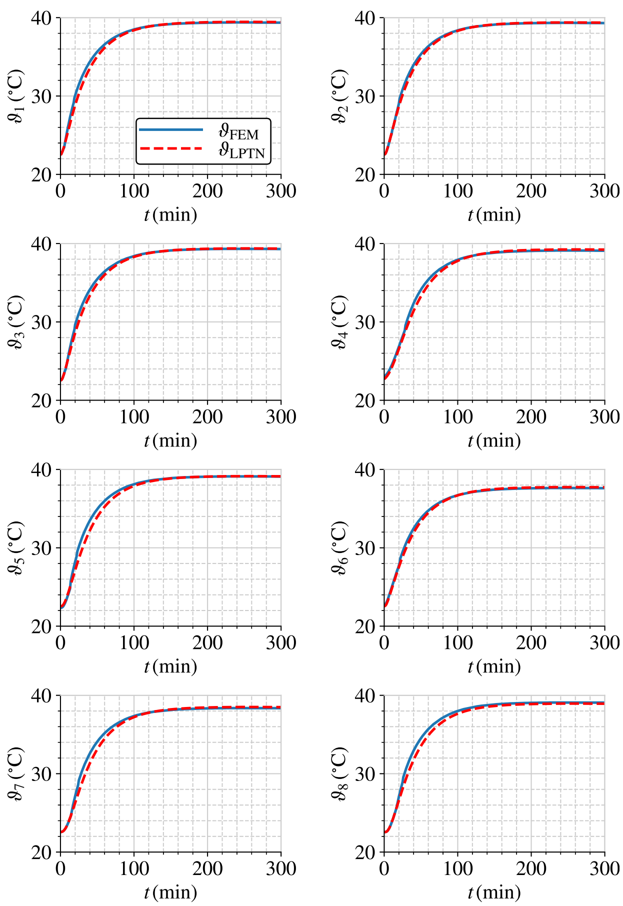


Fig. 7. Average rotor temperatures of the LSPMSM using the LPTN and average temperatures using the FEM in transient state.

The average temperatures of nodes by LPTN are nearly identical to the average temperatures of nodes calculated by FEM. The little difference in average temperatures occurs at node number seven which is situated in the shaft internal air region. The average temperatures differs because in LPTN is considered symmetrical temperature profiles in axial direction, while FEM considered asymmetrical temperature profiles in axial direction [13].

V. CONCLUSION

This paper describes the automated parametric geometry creation using Python programming language to control the FreeCAD modelling software. Then, this automated parametric geometry creation is used on the rotor of the LSPMSM. Next, the FEM thermal analysis the rotor geometry of the LSPMSM is performed. This FEM thermal analysis can be used to validate LPTN. The thermal analysis results of the LSPMSM rotor by FEM and by LPTN analysis are in very good agreement. Automated FEM thermal analysis with automatic geometry generation can be used for various input parameters such as geometric dimensions, convective boundary conditions and heat losses.

REFERENCES

[1] B. Zöhra, M. Akar, and M. Eker, "Design of a novel line start synchronous motor rotor," *Electronics*, vol. 8, no. 1, pp. 1–18, 2019. [Online]. Available: http://www.mdpi.com/2079-9292/8/1/25

[2] A. Boglietti, A. Cavagnino, M. Pastorelli, D. Staton, and A. Vagati, "Thermal analysis of induction and synchronous reluctance motors," *IEEE Transactions on Industry Applications*, vol. 42, no. 3, pp. 675–680, 2006. [Online]. Available: http://ieeexplore.ieee.org/document/1634672/

[3] P. S. Ghahfarokhi, A. Kallaste, A. Belahcen, T. Vaimann, and A. Rassolkin, "Review of thermal analysis of permanent magnet assisted synchronous reluctance machines," in *2016 Electric Power Quality and Supply Reliability (PQ)*. IEEE, 2016, pp. 219–224. [Online]. Available: http://ieeexplore.ieee.org/document/7724116/

[4] A. Boglietti, A. Cavagnino, and D. Staton, "Determination of critical parameters in electrical machine thermal models," *IEEE Transactions on Industry Applications*, vol. 44, no. 4, pp. 1150–1159, 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4578771/

[5] A. EL-Refaie, N. Harris, T. Jahns, and K. Rahman, "Thermal analysis of multibarrier interior pm synchronous machine using lumped parameter model," *IEEE Transactions on Energy Conversion*, vol. 19, no. 2, pp. 303–309, 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1300694/

[6] P. Mellor, D. Roberts, and D. Turner, "Lumped parameter thermal model for electrical machines of tefc design," *IEE Proceedings B Electric Power Applications*, vol. 138, no. 5, pp. 205 – 218, 1991. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/ip-b.1991.0025

[7] Y. A. Cengel and A. J. Ghajar, *Heat and Mass Transfer: Fundamentals and Applications*, 5th ed. New York: McGraw-Hill Education, 2015.

[8] J. Pyrhönen, T. Jokinen, and V. Hrabovcová, *Design of rotating electrical machines*, 2nd ed. Chichester: Wiley, 2014.

[9] *FreeCAD 0.18.3*, 2019. [Online]. Available: https://github.com/FreeCAD/FreeCAD/releases/tag/0.18.3

[10] *Python 3.6.0*, 2016. [Online]. Available: https://www.python.org/downloads/release/python-360/

[11] V. Elistratova, "Optimal design of line-start permanent magnet synchronous motors of high efficiency," Master's thesis, Ecole Centrale de Lille, 2015. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01308575/document

[12] G. D. Demetriades, H. Z. de la Parra, E. Andersson, and H. Olsson, "A real-time thermal model of a permanent-magnet synchronous motor," *IEEE Transactions on Power Electronics*, vol. 25, no. 2, pp. 463–474, 2010. [Online]. Available: https://ieeexplore.ieee.org/document/5204128/

[13] M. Toman, R. Cipin, and P. Vorel, "18th international conference on mechatronics - mechatronika (me)," pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8624804

Overview of the Current State of the Art of Existing Energy Community Concepts in Europe and their Possible Implementation in the Czech Republic

Jan Vojtech, Michal Ptacek
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xvojte11@vut.cz, ptacekm@vut.cz

***Abstract*— This paper presents an overview of the current state of the art of energy communities in Europe and discusses real examples/concepts implemented in some countries that fulfill the issued EU directives by different technical ways. Therefore, it brings so-called allocation keys used in the individual energy community concepts for the distribution of produced energy between consumers and shows the important differences of these keys and their impacts (benefits/disadvantages) for consumers or distribution system operators, so distribution tariffs and the future evaluation of allocation keys are also presented. Furthermore, there is the discussion which energy community concepts (implemented in European countries) or their specific parts can be transferred to the Czech Republic, especially in the context of the current laws and their preparations.**

Keywords—Energy community, renewable energy sources, allocation key, tariff, distribution system

I. Introduction

The idea of energy community (EC), i.e. local energy sharing, can be dated to the second half of the last century and it can be considered that it especially originated in various environmental (ecology) movements in the countries of Western Europe. The main goal of this concept was to reduce greenhouse gas emissions associated with the energy sector. The concept became relevant again in the last ten years after the release of the directives of the European Commission and the Council of the European Union (EU) and after the significant improvement in the availability of renewable energy sources (RES) due to decreasing RES investment costs.

An energy community is a group of people who share or jointly produce electrical energy for their own usage. The establishment of an energy community can be motivated by both environmental thinking and for the purpose of saving electrical energy costs. However, this concept does not represent only sharing electrical energy, but also thermal energy or gas produced in local biogas stations. Above all, in the field of electrical energy, this is a new challenge for distribution system operators (DSOs) due to the possible change in the power flows of electrical energy in distribution system (DS) and due to the additional requirements to measurement, transmission and evaluation of data on the exchange of energy mainly in low voltage DS.

EC was formally defined by two European directives [1] and [2]. The guidelines were created one year apart, and each defines community energy in a different way.

The directive [1] on the promotion of the use of energy from RES, published in 2018, defines EC under the name “renewable energy communities” (REC). A REC could produce, consume, store and above all share electrical and thermal energy. REC would be a legal entity with open access for natural persons, small and medium-sized enterprises and local authorities including municipalities. After joining, it is necessary to preserve the right of members to freely choose an electrical energy supplier, just as in the case of a regular consumer. It is emphasized that the primary purpose of the REC is not to generate profit, but to bring environmental, economic, or social benefits to the members of the community or around operation area. Members of the REC may be partially or completely exempt from certain fees (e.g., for distribution) or taxes. However, the exemption must not occur if the costs are transferred to non-members of the community, or if the public DS is used for sharing.

The directive [2] on common rules for the internal market for electrical energy, published in 2019, defines EC under the name “citizen energy community” (CEC). This directive shares elements of definition, purpose, and rights with [1], but differs slightly in some parts. The important difference is that directive [2] does not talk about electrical and thermal energy, but only about electrical energy. The rights of the CEC are to be expanded to include the provision of energy efficiency services, charging of electric vehicles and other services provided to members of the CEC. It is mentioned that EU member states could allow CEC to own, set up, buy, and lease DS. In such a case, non-members of the community must be able to connect to such a local DS without discriminatory charges for that consumer. The granting of rights associated with the DS is at the discretion of each of the EU member states.

II. Energy Community Models

Existing EC can basically be divided into two models. The first is the EC of an apartment building, the second model is the EC of municipalities, districts or entire regions. The EC of an apartment building is not considered a true community by some countries, as the exchange of energy takes place only in the given building and is therefore not very different from, for example, factories with rooftop installations of photovoltaic panels. But the allocation keys for energy from a joint production plant are used, which is typical for EC. The EC of municipalities, districts and regions can benefit not only from the installation of renewable sources on buildings, but also from stand-alone wind or photovoltaic power plants. Community members can share the costs of investment and operation of resources in the form of their own fees. The exchange of energy usually takes place via the public DS.

III. Allocation Keys

According to [3], allocation keys are used to allocate the produced energy. Allocation keys serve only for the economic (or invoicing) equalization of produced and consumed electrical energy and do not necessarily respect the physical flow of electrical energy. The period for which energies are evaluated may vary depending on state legislation. It is usually evaluated every 15, 30 or 60 minutes. Reference [3] informs that allocation keys are divided into static allocation key (SAK) and dynamic allocation key (DAK), or a combination of both. When using a SAK, production is distributed according to a fixed value. In the evaluated time, one consumer is therefore allocated an agreed share up to the agreed maximum (e.g., 10 kWh), but it must always be less than or equal to the consumption. Energy not allocated to any consumer is sold to an agreed energy trader. The DAK distributes energy proportionally according to the current amount of consumption and thus enables more efficient use of energy from the community source. There can also be a combination of both keys, i.e., re-allocation of energy surpluses remaining after the first allocation, which is allowed in France as is mentioned in [4].

The difference in the principle of SAK and DAK can be seen in Fig. 1 and Fig. 2. In the sample cases the production of 10 kWh every 15 minutes is distributed between two consumers. With a SAK, energy is equally distributed between both consumers. In the second quarter hour with time stamp 30 minutes, consumer B is not using his share and the electrical energy is therefore sold to the energy trader. Consumer A, on the other hand, uses more energy than his share and had to buy electrical energy from an energy trader. This problem is eliminated by the DAK in Fig. 2. In the second quarter of an hour with a time stamp of 30 minutes, electrical energy was redistributed in favour of consumer A. If both consumers have a combined consumption higher than the volume of energy produced, a proportional distribution will occur between them. However, if consumer B had a significantly higher consumption than consumer A, it may happen that only a small part of the energy will be allocated to consumer A. Therefore, it is appropriate to divide the costs between consumers according to how much energy they use from the community source.

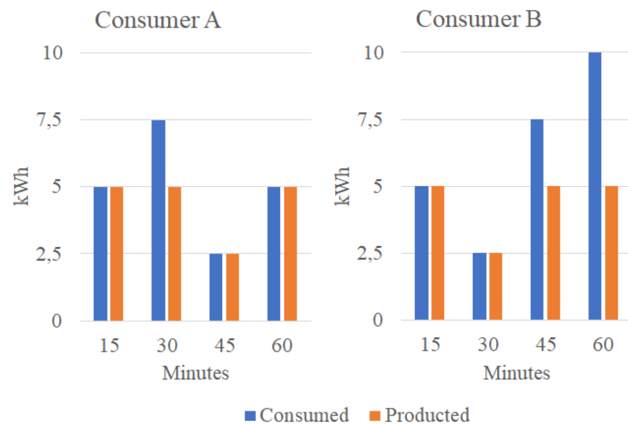


Fig. 1. A demonstration of static allocation key in a case of 10 kWh energy production in every 15 minutes

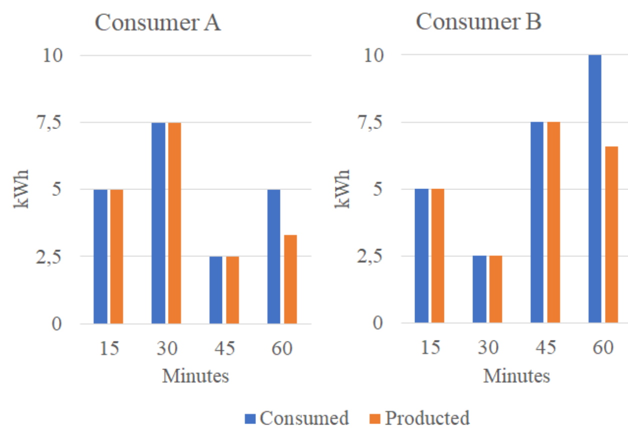


Fig. 2. A demonstration of dynamic allocation key in a case of 10 kWh energy production in every 15 minutes

IV. European Examples

EC are currently not very widespread across Europe. The reason for this is the lack of legislation in many countries preventing community sharing over a public DS. The pioneers in this field are France, Spain, and Austria. Each of the countries approached the EC in a different way.

A. France

Reference [3] mentions that France was the first country to allow community sharing of electrical energy under the name of collective self-consumption. The first definition appeared in the legislation in 2016 and since then there have been adjustments based on experience. Total power of no more than 3 MWp can be installed within the entity of one EC. Reference [3] also mentions that the area of the EC is set so that the farthest places involved in the EC must be within a maximum distance of 2 km in an urban area or 20 km in a rural area. The areas of two EC may overlap, but each place of production or consumption may be involved in at most one of them. The place of production and consumption must be equipped with smart meters and connected to the low voltage DS, although electrical energy can also be transmitted through the medium voltage DS. Energy allocation is evaluated in 30-minute intervals.

According to [4], it is allowed to use SAK or DAK, including their combination. Multiple allocation keys can be selected for different parts of the year. There are no discounts on distribution fees for energy transmitted within the community.

B. Spain

In [3] is mentioned that Spain also calls EC as community self-consumption. Installed power is not limited both production and consumption sites using the public distribution network must be connected in one of the three following ways:

- In the low voltage DS behind one transformer station.
- In the low-voltage DS at a maximum distance of 500 meters from the most distant points.
- In the same cadastral district determined by the first 14 digits.

Reference [5] informs that energy allocation is evaluated in hourly intervals. It is possible to choose between a SAK or SAK with a variable value. No distribution fees are paid for the energy produced and transmitted in the low voltage DS.

C. Austria

Austria was the last of the mentioned countries to enable community sharing. Austria has separately defined REC and CEC in its legislation. The difference between them is in resources, area, rights, and distribution fee discounts.

Federal law [6] informs that CEC can use any sources to produce electrical energy and can operate in the entire area of Austria and offer services to its members (for example charging electric cars). In comparison with the CEC, the REC can only use renewable resources and can only operate at the low and high voltage levels. The REC is further divided into local and regional EC according to the voltage level they connect to. The local EC is only at the low voltage level. The regional REC is at both the low and medium voltage level, alternatively only at the medium voltage level.

Reference [7] mentions that energy exchanged within the REC is discounted according to how much of the grid is used, but energy bought or sold to an energy trader has full distribution fees. The following discounts are applied:

- A low-voltage customer in the local REC receives a 57 % discount on shared energy due to not using medium-voltage and high-voltage DS.
- A medium voltage customer the REC receives a 64 % discount on shared energy, because he only pays for the medium voltage DS and receives a discount for the high voltage DS.
- A low voltage customer in the regional REC receives a 28 % discount on shared energy, as they must pay for the medium voltage DS in addition to the low voltage DS.

Federal law [6] allows a selection between SAK and DAK for both REC and CEC. Evaluation of energy flows based on allocation keys takes place in 15-minute intervals.

V. Situation in the Czech Republic

From the beginning of 2023, community sharing within apartment buildings in the Czech Republic is allowed by [8]. Before the new legislation came into force, it was only possible to combine the entire house under one supply point. However, consumers thereby lost their rights to freely choose an energy supplier.

In the new legislation [8], a leading and a subordinate collection point are introduced. An installed community source is connected to the leading collection point. Only SAK is allowed. The energy distributed according to the allocation key is evaluated in 15-minute intervals. Adequate electrical energy meters must be installed for a 15-minute measurement. It is not necessary to involve all supply point in the apartment building EC.

A. The Future of Energy Act

The amendment to the Energy Act [9] is currently being revised following a comment procedure. In the future, an EC and a REC could be established. Both entities could generate, consume, share, sell electrical energy and provide other services to their members. The difference between them would be that even medium-sized enterprises can participate in the REC in contrast to EC. But the problem with the current proposal is the missing mention of the possibility of building community energy storage and allowing providing flexibility. The proposal does not even mention the possibility of operating a local DS.

B. Data Hub

In summary report [10] from Ministry of Industry and Trade is written that Data Hub is to be established for the purpose of keeping a database of data from electrical energy meters. It would also be possible to find out real-time information about the possibilities of providing flexibility based on the state of the network. It would also be possible to offer and demand flexibility here. The Data Hub would also participate in the coordination of the preparation for the operation of the electrical energy system of the Czech Republic.

VI. Distribution tariffs

Current EC of apartment buildings pay only fixed DS tariff for energy shared within the community, and variable DS tariff depending on the amount of energy consumed are exempt. In the future, when the legislation allows sharing over a public DS, this situation might not be sustainable. It is therefore advisable to think about whether to provide discounts, or how big the discounts should be.

The first option is to provide a discount along the lines of Austria, which may not be sustainable in the long term. Costs for servicing DS would be passed on to consumers without EC membership.

The second possibility is to change the ratio between the fixed DS tariff for the size of the circuit breaker and the variable DS tariff for energy taken. According to unpublished document [11] made by DSO, a large part of the network operation costs are fixed, the distribution system operator would be interested in turning this ratio around and then it would be possible to provide discounts from the variable tariff adequate to the benefits of the community. On the basis of unpublished statistics obtained from the Energy Regulatory Office and on the basis of the price list [12] published by the energy supplier E.ON company, the average payment for distribution tariffs were calculated. Current fixed-to-variable ratio is approximately 30/70. However, calculations also showed that changing ratio to 70/30 would lead to the fact that the customer would pay a large part of the distribution fees in the fixed tariff and would be motivated to reduce consumption primarily by the commercial prices of energy.

The third option is to leave consumers without discounts on DS tariff and motivate only by saving on power electrical energy.

VII. Conclusion and Discussion

The paper shows that EC are and will be gradually implemented in the EU. The reason is the common environmental interest associated with the fulfilment of partial targets for the development of renewable and emission-free sources. ECs are not the only solutions to achieve targets, but they open the other way. The paper also shows that each state approaches EC implementation with different technical and non-technical solutions. Before implementation in the Czech Republic, it is necessary to prepare extensive impact studies and compare the concept that is most suitable for local conditions from the point of view of EC and DSO.

In connection with this, it will be necessary to evaluate the various allocation keys. Anonymized data on the consumption of an example apartment building with 26 electricity meters were obtained for this purpose from EG.D company. Data were recorded every 15 minutes for the entire year 2022. Data from the apartment building will be paired with data from the photovoltaic system. This will also obtain values determining the use of the installed capacity of the photovoltaic system within the apartment building and the potential overflow into the network. After the evaluation at the apartment building level, the community will be expanded to include family houses. The monitored data will result in recommendations for consumers, distribution system operators and possibly also the Energy Regulatory Office.

Acknowledgment

Authors would like to thank the Energy Regulatory Office and EG.D company for their cooperation in providing and processing data.

References

[1] Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources (recast) [online]. Luxembourg: Publications Office of the European Union, 2018 [cit. 2022-10-13]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018L2001&from=EN

[2] Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU (recast) [online]. Luxembourg: Publications Office of the European Union, 2018 [cit. 2022-10-13]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0944&from=EN

[3] M. Ander, P. Novotny and M. Madej. Energetically active consumers in selected EU countries [Energeticky aktivní spotřebitelé ve vybraných zemích EU] [online]. Prague: Ministry of Industry and Trade, 2021 [cit. 2022-12-13]. Available from: https://www.mpo.cz/assets/cz/energetika/vyzkum-a-vyvoj-v-energetice/resene-dokoncene-projekty-a-jejich-vystupy/2021/2/Energeticky-aktivni-spotrebitele-ve-vybranych-zemich-EU.pdf

[4] Collective self-consumption, principle and current state in France in 2021 [Autoconsommation collective, principe et état des lieux en France en 2021. Encyclopédie de l'énergie] [online]. Grenoble: Grenoble INP-Ense3, 2023 [cit. 2023-03-09]. Available from: https://www.encyclopedie-energie.org/autoconsommation-collective-principe-etat-lieux-france-2021/

[5] Frequently asked questions about self-consumption [Preguntas frecuentes sobre autoconsumo] [online]. Madrid: Institute for the Diversification and Saving of Energy, 2022 [cit. 2023-03-09]. Available from: https://www.idae.es/sites/default/files/documentos/idae/tecnologias/energias_renovables/OFICINA-de-AUTOCONSUMO/2022-03-30_99_FAQ.pdf

[6] Federal Act Providing New Rules for the Organisation of the Electricity Sector (Electricity Act 2010 – ElWOG 2010) [online]. Vienna: Federal Ministry of Finance, 2022 [cit. 2022-11-12]. Available from: https://www.ris.bka.gv.at/Dokumente/Erv/ERV_2010_1_110/ERV_2010_1_110.pdf

[7] Lower grid costs for REC. Austrian coordination office for energy communities [Niedrigere Netzkosten für EEGs. Österreichische Koordinationsstelle für Energiegemeinschaften] [online]. Vienna: Klima- und Energiefonds, 2021 [cit. 2023-03-09]. Available from: https://energiegemeinschaften.gv.at/niedrigere-netzkosten-fuer-eegs/

[8] CZECH REPUBLIC. Decree of December 6, 2022 amending Decree No. 408/2015 Coll., on Electricity Market Rules, as amended [Vyhláška ze dne 6. prosince 2022, kterou se mění vyhláška č. 408/2015 Sb., o Pravidlech trhu s elektřinou, ve znění pozdějších předpisů]. In: Collection of laws. Prague: Ministry of the Interior, 2022, year 2022, amount 182, number 404.

[9] Draft law amending Act No. 458/2000 Coll., on business conditions and the performance of state administration in the energy sector and on the amendment of certain laws (Energy Act), as amended, and Act No. 406/2000 Coll., on energy management, as amended [Návrh zákona, kterým se mění zákon č. 458/2000 Sb., o podmínkách podnikání a o výkonu státní správy v energetických odvětvích a o změně některých zákonů (energetický zákon), ve znění pozdějších předpisů, a zákon č. 406/2000 Sb., o hospodaření energií, ve znění pozdějších předpisů] [online]. Prague: Office of the Government of the Czech Republic, 2022 [cit. 2022-12-13]. Available from: https://odok.cz/portal/veklep/material/KORNCKUGPN1S/

[10] Summary report April 2022: Tender No. 23, DataHub [Souhrnná zpráva duben 2022: Zadávací list č.23, DataHub ČR] [online]. Prague: Ministry of Industry and Trade, 2022 [cit. 2022-12-20]. Available from: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2022/6/ZL23_zprava_Q1_22_full_final.pdf

[11] EG.D. Development of energy communities and possible price discounts for distribution [Rozvoj energetických společenství a možné slevy ceny za distribuci]. Brno, 2022.

[12] Price list Complete electricity 36 March 2021 [Ceník Komplet elektřina 36 Březen 2021] [online]. České Budějovice: E.ON, 2021 [cit. 2023-03-12]. Available from: https://www.eon.cz/getmedia/0931a609-bfc2-414d-b475-8b0fe3ea22e9/cenik-komplet-elektrina-36-brezen-2021-distribucni-uzemi-egd.pdf?ext=.pdf

Providing flexibility with detachable loads Procurement and commodity product prediction

Jan Sykora, Lukas Radil
Department of Electrical Power Engineering
Brno University of Technology
Brno, Czech republic
jan.sykora2@vut.cz, radil@feec.vutbr.cz

Martin Pilař
Department of Procurement and commodity product prediction
ČEZ Prodej, a.s.
Praha, Czech republic
martin.pilar02@cez.cz

***Abstract*— Flexibility is generally considered as the potential of a facility (generation/consumption or storage) to actively change the amount of its generation/consumption/accumulated energy based on price signals or direction. This control strategies is managed to correct deviations of the electricity grid or to ensure that the purchase/sale of electricity is profitable for the flexibility provider. This paper presents a comparison of the control of heat pump consumption to provide consumption flexibility.**

Keywords— flexibility, heat pump, demand response, baseline, rebound effect, aggregator, renewable energy, national action plan, smart grid

I. Introduction

Currently, several working groups in Czech Republic are dealing with the concept of flexibility, for example, we can mention the working group of the AKU-BAT Association, the ČK Cired working group, or the DFLEX project led by ČEPS. In general, flexibility can be seen as a newer approach with respect to ripple control technique (HDO), which is still technically capable of working, but no longer offers higher utility value to the requirements of electricity traders. Flexibility is now also associated with the concepts of Peak Shaving and Load Shifting. The first option reduces consumption during peak periods by limiting certain types of loads in order to avoid unwanted energy peaks. However, the conditions for the possibility of using such an intervention need to be evaluated before the actual operation starts so that the flexibility provider is not affected.

The second possibility of using flexible management is load shifting, which is shifting the load to a part of the day that is more beneficial for both the provider and the system operator.

Flexibility can also be found, among other things, in households, especially those with electric heating systems. Due to the thermal capacity of the household, it is possible to heat the room space to the required temperature during off-peak periods or to shut down the pump for a period of time during peak periods, but with respect to the comfort of the provider. One of the barriers to this application of controlled flexibility may be the unwillingness of customers to provide flexibility, so in the future it is necessary to come up with a suitable direct or indirect way to motivate the customer.

II. Terms related to flexibility

Prosumer - is a customer who has an electricity generation plant connected at his point of connection to cover his own consumption with the possibility to supply part of this electricity to the electricity grid.

Demand side response (DSR) - represents a change in customer demand or supply in response to a price signal or command.

Flexibility Provider (PoFl) - This is an entity providing flexibility individually or through an aggregator.

Flexibility Resource - this is a facility providing flexibility by consuming, storing or generating electricity.

Activation of a flexibility resource - this is an action by the aggregator or flexibility provider to reduce or increase consumption or generation or to change the output of the flexibility resource.

Positive flexibility potential - this is a potential reduction in load on the consumption side or a potential increase of output on the production side [2].

Negative flexibility potential - this is a potential increase in load on the consumption side or a potential decrease of output on the production side [2].

Flexibility Purchase - the conclusion of a flexibility contract between the aggregator and the flexibility provider, which allows the aggregator to vary the amount of electricity consumed and supplied at the provider's point of consumption. This contract should include the parameters of the flexibility provision (maximum load change, duration) and the method of financial compensations.

Baseline - this is a diagram of the estimated value of consumption or generation that would occur without the activation of flexibility [3].

Rebound effect - this is the potential shift in the flexibility provider's consumption or production over time due to the activation of the flexibility, which causes further deviation of the entity responsible for the flexibility provider's deviation.

Baseload - this is the minimum continuous consumption of electricity. Its delivery zone is every day of the week from 0:00 to 24:00 [4][6].

Peakload - this is the load during the peak demand for electricity. Its delivery zone is every weekday from 08:00 to 20:00 [6].

Offpeak load - this is electricity off-peak. Its delivery zone is every weekday from 20:00 to 08:00 the following day [6].

III. Methodologies for Establishing the Baseline

The default load diagram relates to a specific point of consumption and represents the magnitude of consumption over time, assuming that flexibility would not be activated, Fig. 1. This diagram is therefore a basic tool for evaluating the potential and actual flexibility of a given provider, especially the impact on the trader.

The baseline, as the default load diagram can otherwise be called, allows the measurement of the load on the customer sites and its good design benefits all stakeholders by aligning the activation conditions and interests of flexibility providers, aggregators, suppliers and network operators. However, the baseline is still only an estimate and the choice of an appropriate methodology is a prerequisite for its proper design.

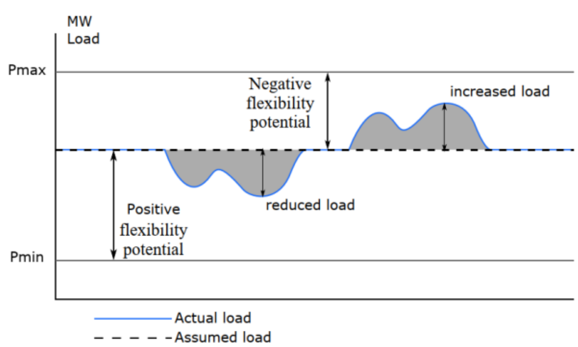


Figure 1 Load flow and load changes

There are several methodologies that can be applied, each bringing different benefits and different possibilities of application. However, 3 key factors are critical to its selection - accuracy, simplicity and integrity.

A. *Baseline type 1*

Today, the most common methodology is the baseline type 1, with common variations including:

1) Averaging,
2) regression,
3) moving average,
4) comparable day.

The characteristic elements of a Type 1 baseline are the shape of the baseline based on average historical load, measured data for each sampling point, the use of data measured on the days immediately prior to activation, and in addition, it may also consider the effects of weather or other events. Another major advantage of this method is its relative ease of application and, with the appropriate choice of time horizon and methodology to exclude inappropriate data, its relative accuracy [1].

B. *Baseline type 2*

Unlike the other methods, the Type 2 methodology does not use data measured at specific consumption points, but works with aggregated data from multiple customers with similar consumption and behaviour. It then uses the measurements at several individual sites to produce an estimate of the average load of the units in question and then uses this data to redistribute it between the individual consumption sites. This method is only advantageous in the savings for reducing the number of metering devices.

C. *"Before and after" method*

The before/after measurement method can be applied especially when the activation period is of short duration, typically from 10 minutes to a maximum of 3 hours. In the case of activation of longer duration, it can no longer be considered reliable.

D. *Reference group methodology*

The reference group method, as the name suggests, uses a reference group of similar providers to determine a substitutional baseline.

E. *Rebound effect*

In the case of DSR, the activation of flexibility also has an impact in the time following the activation, a phenomenon we refer to as the "Rebound effect". Rebound effect arises as an impact of postponed consumption. The rebound effect also has an impact on the trader as it prolongs the duration of the deviation from the established type diagram on which the trader purchases energy, and therefore, in the case of providing flexibility through an independent aggregator, it is necessary to determine how to settle this effect [5].

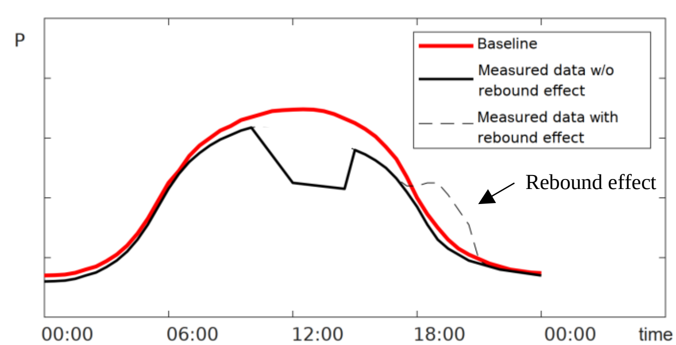


Figure 2 Rebound effect

IV. Analysis and Evaluation of Measured Data

In the framework of cooperation with ČEZ Prodej, a.s., data from real measurements were obtained, which took place from December 2021 to March 2022. The data from this dataset usable for subsequent evaluation amounted to only 20 controlled consumption points and 6 uncontrolled ones in the period 13 December 2021-30 January 2022. The objects examined in this work fall under tariff rate D57d, which is intended for consumers using an electric heating appliance as the main source of heating, especially a heat pump. It is a two-tariff rate, where a low tariff is used for 20 hours per day.

The data evaluation was based on a modified Baseline Type1 methodology with temperature corrections.

However, the size of the aggregation intervals significantly effects the duration of the rebound effect, the duration of which is evaluated by overcoming the initial load diagram by the measured load in the direction in which the activation acted. Throughout the measurement, it was possible to observe states where this limit was not reached due to external influences. These formally uncompleted rebound effect waveforms were subsequently plotted at 120 minutes. As can be observed in the waveforms (Fig. 3,4,5 and Fig. 6) the influence of the aggregation intervals is noticeable. The load change waveform for a aggregation interval of 15 min indicates a rebound effect duration on average up to 45 minutes longer than what can be achieved using a aggregation interval of 10 min. Fig. 2 through Fig. 5 show the rebound effect response for aggregation intervals of 1 min, 5 min, 10 min, and 15 min.

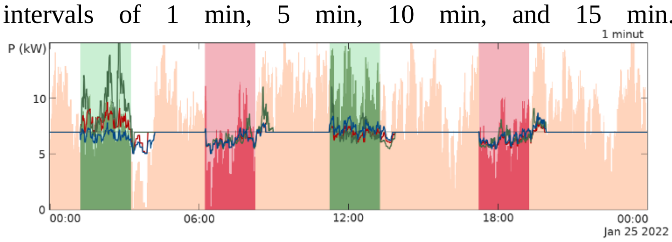


Figure 3 Relative load change for aggregation interval 1 minute

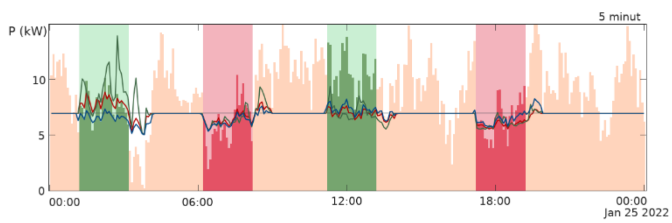


Figure 4 Relative load change for aggregation interval 5 minute

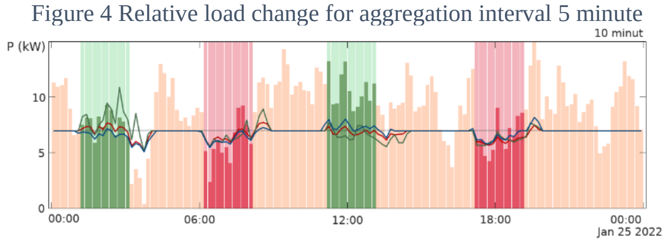


Figure 5 Relative load change for aggregation interval 10 minute

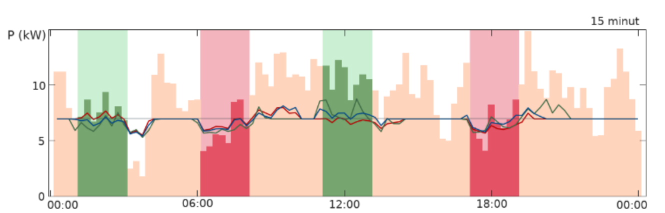


Figure 6 Relative load change for aggregation interval 15 minute

However, if the activation of flexibility is based on commands issued by an independent aggregator, it distorts the expected load diagram of the point of consumption and thus significantly affects the trader. Thus, for subsequent settlement, a substitute baseline must be established based on the selected methodology. This may be based, for example, on historical data or data obtained from reference groups. A correction is then made, which may be based on pre- and post-activation measurements, or weather or other possible variables. From the subsequent evaluation of the measured load and the established substitutional baseline, an evaluation of the impact of activation on consumption at the time of activation can be made.

V. Frequency of Heat Pump Response to an Command

The average frequency of response to an command is the basic information indicating the success of the activation. For this statistic, the data measured in the spring period 2021 from 10.03.2021 to 30.06.2021 can be used, since this file contains information about the issued command in addition to information about the actual execution. The group used in this measurement consists of 6 examples of customers providing flexibility from the same location, identified as shown in Table 1. For this evaluation, I do not consider the instruction as a whole, but as each separate minute that the issuance/execution of the command lasted.

Tab. 1 commands to reduce heat pump power input

command	executed	issued
device_001	2902	3180
device_002	1430	2041
device_003	8660	8762
device_004	2616	3359
device_005	2741	3240
device_006	3205	3660
Total	21554	24242

Command success rate is 88.91%

Tab. 2 commands to increase heat pump power input

command	executed	issued
device_001	1098	3005
device_002	2892	3614
device_003	480	780
device_004	1199	2496
device_005	1340	3067
device_006	1851	9059
Total	8860	22021

Command success rate is 40.23%

Measurements show that the instruction to increase the power has less than half the success rate of the instruction to decrease. This is mainly due to the nature of the heat pump, which tends to maintain the temperature at the upper limit of the comfort limit and switch on when the limit drops. Its control override temperature loop. Consequently, there is considerably less scope for increasing the input power than for decreasing it. This fact is well noticeable in the waveforms from the whole measurement, where the boost instructions are significantly more divergent and less effective.

VI. Conclusion

The growing share of volatile resources in the distribution systems brings risks associated with grid instability. One possible solution is to use flexibility on the consumption side. For the residential sector, the use of electric heating management represents the highest potential for this purpose.

In the context of designing appropriate methodologies for determining the initial load diagram, two of the three methodologies tested produced satisfactory results. There was a significant intersection between these methodologies, with the calculated data varying with an average deviation of 1.14% across the project. The low level of variation between these methodologies is the evidence of functionality and reliability of both of them, with even a relatively simple methodology such as the before/after measurement methodology producing satisfactory results. For the reference group methodology, the conclusion is unsatisfactory as it achieved an average deviation of 14.78% when compared to the primary baseline type 1 methodology. This level of bias, together with the observed patterns from the entire measurement, led to the conclusion that the methodology was unreliable. This is mainly due to the variability of the individual objects. Within a small portfolio size, a baseline cannot be established with sufficient precision based on a different group of customers. This is mainly due to the number of variables, the differences between objects and the randomness in user behaviour.

Within the commands themselves, there was a significant difference between the increase and decrease commands. From the measured data, the success rate of each type within the frequency of response to the instruction was found to be 88.91% for the decrease instruction and 40.23% for the increase instruction. A significant disparity between these values can be observed, this is mainly due to the nature of the heating system which tends to keep the temperature in the building at the upper end of the comfort limit, therefore not providing the necessary scope for negative flexibility. This finding is consistent with the assumptions and is also supported by the data from the evaluation of the amount of flexibility gained, where the amount of negative flexibility was generally lower and more volatile compared to the amount of positive flexibility.

An interesting outcome is the finding of the effect of the daytime of a given instruction, with increasing instructions typically achieving higher efficiency during nighttime hours and decreasing instructions achieving the opposite during afternoon hours. this is primarily due to the typical temperature profile of residential properties during the day. This behaviour is favourable due to the orientation of consumption peaks during the day.

In terms of the observation of the duration of the rebound effect, the work did not meet expectations, mainly due to the uniqueness of the individual results, which led to significant variation and inaccurate evaluation of the data obtained. The work then made observations about the impact of the aggregation interval on this parameter.

Based on the evaluated data, it is therefore possible to proceed with the recommendation to use significantly lower aggregation intervals such as 5 or 10 minutes. This aggregation interval will ensure higher accuracy and sensitivity just for the evaluation of the impact on deferred consumption.

The main contributions of the paper includes an introduction to the issue, a suggestion of possible methodologies for establishing a baseline load diagram, and an insight into the behaviour of the small portfolio of households with a heat pump used in the flexibility framework, including illustrations of different types commands and aggregation intervals.

References

[1] DELOITTE, Role agregátora v české energetice [online]. Praha, 2018, , 150s [cit. 2022-01-05]. Dostupné z: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2020/1/

[2] MPO. Závěrečná zpráva I. etapy projektu ZL6 NAP SG [online]. Praha, 2021 [cit.2022-12-27]. Dostupné z: 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

[3] ČERNÝ, Alexandr. Agregace flexibility a využitelnost AMM: Dosavadní zkušenosti z NAP SG ZL6 [online]. 2021 [cit. 2022-12-27]. Dostupné z: https://www.mpo.cz/assets/cz/energetika/vyzkum-a-vyvoj-v-energetice/workshopy/2021/7/Blok-1--prezentace-4---Agregace-flexibility-a-vyuzitelnost-AMM--dosavadni-zkusenosti-z-NAP-SG-ZL6.pdf

[4] HORKÝ, Josef. Jak se cení elektřina [online]. České Budějovice, 3. 5. 2022 [cit. 2022-12-27]. Dostupné z: https://www.eon.cz/byznys-energie/jak-se-ceni-elektrina/

[5] MPO, Závěrečná zpráva I. etapy projektu ZL6 NAP SG:Za rok 2020/2021. NAP SG ZL6 [online]. 30.4.2021, 1-79 [cit. 2022-01-05]. Dostupné z: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2021/5/Zaverecna-zprava-1-etapy-ZL6-MPO.pdf

[6] PXE. Informace o obchodování a produktech [online]. [cit. 2022-12-27]. Dostupné z: https://pxe.cz/cs/derivatovy-trh/informace-o-obchodovani-a-produktech

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.

NASKENUJ MĚ

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

Deployment of Threshold Signatures for Securing Bitcoin Transactions

1st Minh Tran
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xtranm00@vutbr.cz

2nd Petr Dzurenda
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0002-4366-3950

***Abstract*—Blockchain technology, especially Bitcoin, has revolutionized how we think about and manage financial transactions. However, with the increasing demand and usage of blockchain technology, the security of cryptocurrency wallets has become a critical concern. Threshold signatures offer a promising solution to this problem, allowing multiple parties to sign a transaction without revealing their private keys. This article presents an Android mobile Bitcoin wallet application that uses Schnorr-based threshold signatures. The application also deploys smartwatch integration for enhanced security and usability. This integration provides an additional layer of security by requiring physical confirmation from the user before approving any transaction. Our implementation provides a secure and efficient platform for managing Bitcoin assets using threshold signatures while also providing an intuitive and easy-to-use interface for interacting with the application.**

***Index Terms*—Threshold Signature, Secret Sharing, Security of Transactions, Blockchain, Bitcoin, Android, Smartwatch**

I. Introduction

Blockchain technology is a digital ledger that enables transactions to be recorded and verified without the need for a centralized authority or intermediary. This means that transactions can occur directly between individuals or entities without the need for a trusted third party to oversee or validate the transaction. One of the most well-known applications of blockchain technology is Bitcoin [1], a decentralized digital currency operating on a blockchain network. While blockchain offers benefits such as transparency, reduced transaction costs, and enhanced security, it also poses unique security challenges. One such challenge is the use of single-key transactions, which can make them vulnerable to theft or loss.

To address this issue, multiparty cryptographic schemes and threshold signatures can enhance the security and privacy of blockchain-based systems by allowing multiple parties to sign a transaction without revealing their private keys. However, the use of threshold signatures in blockchain-based systems also poses some challenges. For example, there is a need to establish a secure and reliable protocol for distributing and managing the threshold keys. The usage of distributed signature is not a new discovery. Several blockchains, such as Bitcoin, Cardano, and Algorand, already support multi-signature transactions through multi-signature wallets, but this approach has privacy and scalability concerns. All owners and signers are publicly known, and the transaction size increases with the number of signers.

A. Contribution and Paper Structure

This paper presents several contributions: 1) development of a user-friendly Android Bitcoin wallet application that supports both single-key and threshold signatures, ensuring secure storing of private keys and sending of Bitcoins, 2) optimization of the threshold signature scheme introduced by [2], 3) integration of smartwatches into the threshold signature scheme to provide an additional layer of security by acting as hardware wallets, 4) experimenting and investigating the viability of threshold signatures on the Bitcoin blockchain, and 5) conducting extensive testing of the proposed solution.

The paper is organized as follows. Section II outlines the used terminology. Section III presents our design and development of a threshold signature wallet. Section IV presents our experimental results. In the last section, we conclude this work.

II. PRELIMINARIES

In this section, we introduce blockchain and threshold signatures technologies as they are fundamental building blocks of our implementation.

A. Bitcoin blockchain

The purpose of the Bitcoin blockchain is to provide a secure and decentralized way for individuals and organizations to conduct transactions without the need for a central authority or intermediary. The ledger is maintained by a network of nodes that work together to validate transactions and add them to the blockchain. This means that no single entity has control over the network, and all transactions are transparent and visible to anyone who wishes to view them. Transactions are secured using cryptographic algorithms and a consensus mechanism known as proof of work. This ensures that transactions cannot be tampered with or reversed, and that the network remains secure and resistant to attacks. The Bitcoin blockchain is also designed to be immutable. This ensures the integrity of the network and provides a high degree of trust in the system.

The security of Bitcoin transactions relies on the cryptography of private and public keys. Private keys are used to prove ownership of the cryptocurrency. They are needed for signing transactions. They are kept secret by the owner of a Bitcoin wallet. Public keys, on the other hand, are used to verify transactions. They are derived from the private keys using child key derivation and are publicly visible on the blockchain.

From the Taproot hard fork [3], Bitcoin uses the Schnorr signature algorithm for securing transactions. This scheme allows more efficient and secure transactions. Unlike the traditional Elliptic Curve Digital Signature Algorithm (ECDSA) used in Bitcoin, Schnorr signatures allow for multiple inputs to be signed together, reducing the size and complexity of transactions on the blockchain. Accordingly, this permits the implementation of multi-party computation schemes like threshold signature algorithms.

B. Threshold Signatures

Threshold signatures allow multiple parties to sign a document or any other data with a shared secret key without revealing the key itself. This is achieved by dividing the secret key into several parts and distributing them among the parties. A predetermined number of parties must collaborate to generate a threshold signature. This method significantly reduces the risk of the key being lost or stolen and can adapt to changing requirements over time.

Our proposal is based on Ricci et all [2] scheme that involves 3 cryptographic primitives: Schnorr signature, Shamir secret sharing scheme and Paillier cryptosystem. The scheme consists of 3 stages: setup, signing and verification. Signing entities can be divided into main and secondary signers. The main signer initiates the signing process and has the final signature. The secondary signer responds to a request from the main device and has only a partial signature. In Algorithm 1, the Setup stage is described. Security of transmitted data and the ability to compute over encrypted data are achieved by the Paillier cryptosystem and its homomorphic properties.

Algorithm 1 `Setup`

Phase 1 - generation of public and private values

generate at random $d_1^{(j)}, \dots, d_{t-1}^{(j)}$
generate the Paillier's key pair $(pk_{p,j}, sk_{p,j})$
generate at random k_j in $\mathbb{Z}_{q_{EC}}$
compute $pk_j = g^{k_j}$

Phase 2 - computation of Shamir secret sharing values

D_h generates random value $r_{j,h}$ and computes $e_{j,h} = Enc_{pk_{p,j}}(\alpha_j, r_h)$
D_h generates random value $v_{j,h}$ and computes $c_h = e_{j,h}^{\alpha_j^{t-2} * d_{t-1}^h} * e_{j,h}^{\alpha_j^{t-3} * d_{t-2}^h} * \cdots * e_{j,h}^{d_1^h} * Enc_{pk_{p,j}}(k_j, v_{j,h})$
if $h = j$, then D_j computes $f(\alpha_j) = Dec_{sk_{p,j}}(c_j - 1) + d_{t-1}^j * \alpha_j^{t-1} + \cdots + d_1^j * \alpha_j + k_j$

Figure 1 depicts a sketch of the scheme. In the Signing stage, the first thing is to specify who will join the signature. The Signing stage also has two phases. In the first phase, every signer who was specified to join the signature computes a session key. This is done through interpolation. In the second phase of this stage, every signer computes a commitment and sends it to the main signer, who adds all the commitments together and sends the aggregated commitments to secondary devices. After that, every signer computes their partial signature with the Schnorr signature algorithm and sends it to the main signer. The main signer finally aggregates all of the partial signatures into a final signature.

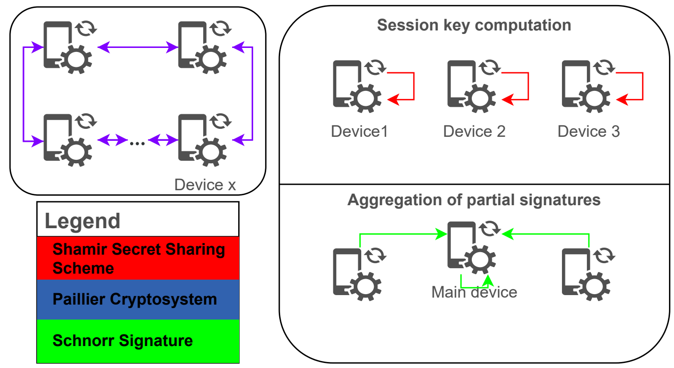


Fig. 1. Stages of the base threshold scheme.

The final Verification stage is done via the Bitcoin blockchain itself. This stage is described in Algorithm 2.

Algorithm 2 `Verification`

compute $e = taggedHash("BIP0340/challenge", r_{sig} + pk + msg)\%n$
compute $R = G * s_{sig} + pk * (n - e)$
if $(R == r)$ **return true** else **return false**

III. DESIGN AND DEVELOPMENT OF THRESHOLD SIGNATURE WALLET

In this section, we introduce the system architecture of our threshold signature wallet application. Then, we discuss our choice of Bitcoin blockchain over other blockchains, present our modification of the base threshold signature scheme, and finally, describe our wallet's implementation details.

A. SYSTEM ARCHITECTURE

The architecture of the wallet application consists of four main components: 1) the wallet back end with communication with the Bitcoin blockchain, 2) the threshold signature module, 3) communication between devices via Bluetooth, and 4) the Graphical User Interface (GUI). The wallet back end is responsible for managing the user's Bitcoin funds and communicating with the Bitcoin blockchain. It consists of several sub-components, including the transaction manager, the key manager, and the blockchain client. The transaction manager is responsible for creating and signing Bitcoin transactions, while the key manager is responsible for generating, storing, and managing the user's private keys. The blockchain client communicates with the Bitcoin network to retrieve transaction data and submit new transactions. The threshold signature module consists of several functions for generating private and public values, computation and aggregation of partial signatures, and verifying the final signature. The validity of the threshold signature is always checked before it is submitted to the Bitcoin network.

Communication channels between our system's components are depicted in Figure 2. Communication between devices is achieved by using Bluetooth technology. The devices must be paired before they can communicate with each other. Once paired, the devices can exchange data using Bluetooth to generate threshold signatures.

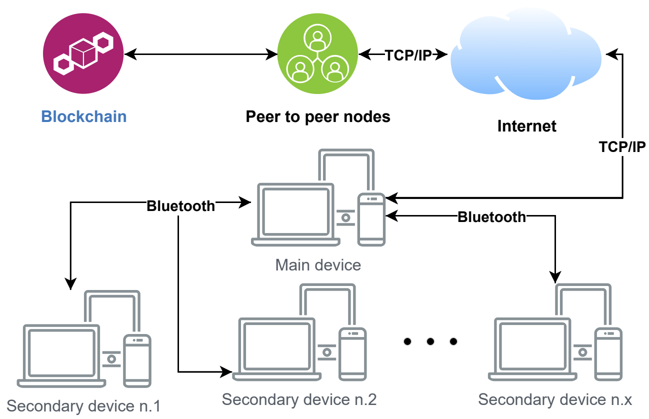


Fig. 2. Communication channels in our system.

The graphical user interface is the front end of the application, which provides a user-friendly interface for the user to interact with the application. It includes several screens, including the wallet screen, the transaction screen, and the settings screen. The wallet screen displays the user's Bitcoin balance, while the transaction screen allows the user to create and sign new transactions. The settings screen enables the user to configure the threshold signature module.

B. SELECTION OF BLOCKCHAIN TECHNOLOGY

There are several factors that lead us to choose the Bitcoin blockchain as the target platform for our threshold signature implementation. One of the most important factors is the fact that the Bitcoin blockchain supports the Schnorr signature scheme, which is a key component of the threshold signature scheme our work is based on. In addition to its support for the Schnorr signature scheme, the Bitcoin blockchain is the most popular today. This means that a wealth of documentation and resources is available for developers working with the platform, making it easier to build and maintain a high-quality implementation. Furthermore, the Bitcoin blockchain is standardized, which means a clear set of rules and specifications must be followed to ensure compatibility with the network. This makes developing and maintaining a secure and reliable implementation easier, as the rules and specifications provide a clear roadmap for developers to follow. To compare available blockchains, some of the most popular blockchains are listed and sorted by their market cap in Table I.

TABLE I
BLOCKCHAINS AND THEIR SIGNING PROTOCOLS

Blockchain	Signing protocol
Bitcoin	Schnorr signature
Ethereum	ECDSA
Binance chain	ECDSA
XRP	ECDSA
Cardano	EdDSA
Polygon	ECDSA
Solana	EdDSA
Polkadot	Schnorr signature

C. THRESHOLD SIGNATURE SCHEME MODIFICATION

The base threshold signature scheme we used in our implementation was initially developed as a standalone protocol and did not consider the specific requirements and characteristics of the Bitcoin blockchain. As a result, we needed to modify the scheme to ensure it was optimized for use in the Bitcoin blockchain. More precisely, the very last stage of the scheme, where partial Schnorr signatures are computed, had to be modified.

In the former scheme, the commitment c is computed as g^r, where r is a random number. In Bitcoin, the value r is not entirely random. Firstly value t is computed as byte-wise xor of the user's private key and a tagged hash of a random number. The value c is then computed as a tagged hash of a concatenation of t, the user's public key and the message. The challenge e is also computed differently as a tagged hash of a concatenation of c, the user's public key and the message. The prove s is computed the same way in both schemes. The final signature is a pair of c and s.

Another problem of the former scheme is the rogue-key attack. This attack can be mitigated by requiring that users prove possession of the secret key, e.g., by attaching a zero-knowledge proof of knowledge to their public keys. We used the solution from MuSig protocol [4] to exercise this approach. The challenge e is computed as $e = H_{agg}(L, X_i) * H_{sig}(\tilde{X}, R, m)$, where $\tilde{X}$ is the aggregate public key corresponding to the multiset of public keys $L = \{X_1, \ldots, X_n\}$.

In Algorithm 3, the signature generation of the base threshold signature is described. The parts highlighted in red are parts that we modified.

D. IMPLEMENTATION DETAILS

The application was developed in Kotlin using Android Studio, with a minimal version of applications programming interface (API) 21: Android 5.1 Lollipop and target API 31: Android 12 Snow Cone. This satisfies about 99.2 % of all Android devices and satisfies the condition for releasing the application on Google Play. The application supports both mobile phones and smartwatches.

The wallet back end is powered by the bitcoin-kpm library [5], which provides a comprehensive set of functions for managing the derivation of Bitcoin keys and creating and signing transactions. We use traditional mnemonic codes with a base of 24 words for the seed. But the users can choose

Algorithm 3 $\texttt{SignatureGen}(\{s_j\}_{j\in\mathcal{J}_t}, m)$

for $j \in \mathcal{J}_t$ do: $\triangleright$ run privately by each D_j (i.e., MD and SDs)
    $t_j = k_j \text{ xor } taggedHash("BIP0340/aux", aux)$
    $c_j = taggedHash(t_j||pk_j||m) \bmod q_{EC}$ $\triangleright$ SDs send c_j to MD
end for
$c = \prod_{j\in\mathcal{J}_t} c_j$ $\triangleright$ run by MD, c and m sent to SDs
for $j \in \mathcal{J}_t$ do: $\triangleright$ run privately by each D_j (i.e., MD and SDs)
    $e = taggedHash("BIP0340/challenge", c||pk||m)$
    $z_j = r_j - e * s_j \bmod q_{EC}$ $\triangleright$ SDs send z_j to MD
end for
$z \leftarrow \sum_{j\in\mathcal{J}_t} z_j \bmod q_{EC}$ $\triangleright$ run by MD
return $\sigma = (c, z)$

their own 25th word to enhance the security even more. All the data is stored in MODE_PRIVATE, so no one except the application has the right to access it. For the blockchain client, we created a library that uses the Blockstream API [6] and communicates with their public node. This library provides a simple and efficient way to interact with the Bitcoin blockchain via posting and getting requests, allowing users to view their balances and perform transactions. The threshold signature module is a critical component of our application, and we developed two libraries to support it. The first library is the Paillier Cryptosystem, which is used to encrypt the secret values before they are distributed to the participants. The second library is the Bitcoin elliptic curve (secp256k1) library, which is used for every computation with the elliptic curve. These two libraries are combined to create the threshold signature module. Furthermore, we implemented Bluetooth connectivity using Wearable Data Layer API to facilitate communication between devices. This allows users to securely communicate and exchange information between devices without needing a network connection or other external infrastructure. Finally, the graphical user interface of our application uses standard .xml and activity architecture. For the main layout, the constraint layout was chosen.

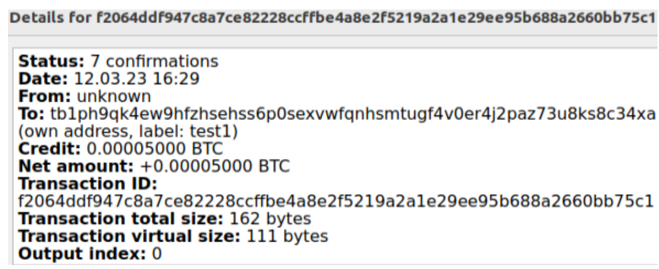
IV. EXPERIMENTAL RESULTS

For our testing scenario, we used a combination of a mobile phone (Samsung Galaxy S9+) and a smartwatch (Samsung Galaxy Watch5 Pro) since they are the most common usage for potential everyday users of our application. This makes our testing scenario a 2-of-2 threshold signature. In order to evaluate the performance of our proposal, we conduct tests on both devices, specifically focusing on the execution time of the setup and signing phases. The findings about the execution time of the scheme and data transfer of the phases for each device are presented in Table II.

Our application successfully builds, signs and transmits transactions which are accepted by the Bitcoin blockchain i.e., the final verification part was also a success. One of those transactions can be viewed in Figure 3, showing a screenshot from the Bitcoin Core application. Even though our threshold signature is constructed using multiple partial signatures, it has the same length and appearance as a normal simple single-key Schnorr signature, making it indistinguishable from any other Schnorr signature. This ensures efficient use of memory space and grants users a high level of anonymity.

TABLE II
BENCHMARKS IN SECONDS OF THE SETUP AND SIGNING PHASES

	Setup		Signing	
	Scheme	**Data transfer**	**Scheme**	**Data transfer**
Phone	0.13	1.04	0.12	0.71
Smartwatch	0.28		0.22	
Total	**1.45**		**1.05**	


Details for f2064ddf947c8a7ce82228ccffbe4a8e2f5219a2a1e29ee95b688a2660bb75c1

Status: 7 confirmations
Date: 12.03.23 16:29
From: unknown
To: tb1ph9qk4ew9hfzhsehss6p0sexvwfqnhsmtugf4v0er4j2paz73u8ks8c34xa (own address, label: test1)
Credit: 0.00005000 BTC
Net amount: +0.00005000 BTC
Transaction ID: f2064ddf947c8a7ce82228ccffbe4a8e2f5219a2a1e29ee95b688a2660bb75c1
Transaction total size: 162 bytes
Transaction virtual size: 111 bytes
Output index: 0

Fig. 3. Transaction made by our application.

V. CONCLUSION

In this paper, we design and implement an Android Bitcoin wallet that supports threshold signatures. The wallet application shows the practicality of using threshold signatures and smartwatches to enhance the security of Bitcoin transactions. The integration of smartwatches as an additional security layer provides users with physical authentication for transaction authorization, making transactions more secure. We have tested the application extensively, and it is capable of handling both single-signature and threshold-signature transactions efficiently and securely. Further improvements can be made to optimize the threshold signature protocol, enhance the user interface, and support other blockchain networks.

REFERENCES

[1] Satoshi Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008 https://bitcoin.org/bitcoin.pdf
[2] Ricci, S., Dzurenda, P., Casanova-Marqués, R., Cika, P. (2022). Threshold Signature for Privacy-Preserving Blockchain. In: , et al. Business Process Management: Blockchain, Robotic Process Automation, and Central and Eastern Europe Forum. BPM 2022. Lecture Notes in Business Information Processing, vol 459. Springer, Cham. https://doi.org/10.1007/978-3-031-16168-1_7
[3] Pieter Wuille, Jonas Nick and Tim Ruffing. Schnorr Signatures for secp256k1. 2020. https://github.com/bitcoin/bips/blob/master/bip-0340.mediawiki
[4] Gregory Maxwell, Andrew Poelstra, Yannick Seurin and Pieter Wuille. Simple Schnorr Multi-Signatures with Applications to Bitcoin. 2018. https://eprint.iacr.org/2018/068
[5] Fabrice Drouin, Bastien Teinturier, Salomon Brys, Richard Myers and Romain Boisselle. Kotlin Multiplatform Bitcoin Library. 2022. https://github.com/ACINQ/bitcoin-kmp/
[6] Nadav Ivgi, Matthew Haywood, Dimitris Tsapakidis, Benthe Carman and Hyunhum Cho. Esplora HTTP API. 2022. https://github.com/Blockstream/esplora/blob/master/API.md

Multi-Level Approach to Cybersecurity Education

1st Willi Lazarov
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czechia
lazarov@vut.cz

2nd Tomas Stodulka
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czechia
xstodu07@vut.cz

***Abstract*—This paper addresses challenges in building a cyber range platform and presents the design, implementation, and testing of a software platform for cybersecurity education, which scope ranges from technical universities to high schools. Our contribution is a multi-level approach to cybersecurity education, which is applied through extensive testing with students from two different technical universities and high schools.**

***Index Terms*—BUTCA, Capture the Flag, Cyber Arena, cyber range, cybersecurity, education, testing, training**

I. Introduction

From the cybersecurity perspective, information and communication technologies are facing numerous types of threats, including malware, password attacks, denial of service (DoS), spam, injection attacks, and many others. These threats have been rapidly increasing recently and are expected to continue to grow at the same or even higher rates. No one is 100% immune to cyberattacks, from government organizations to smaller family-owned businesses or ordinary internet users. Therefore, it is crucial to educate both the public and experts so that they can identify potential attacks, prepare for their consequences, and generally know how to mitigate these security threats. These requirements can be fulfilled by a cyber range (CR) [1], [2] – the platform responsible for the complete preparation of simulated scenarios where users are exposed to different challenges. Cyber ranges are innovative solutions designed to prepare security environments for effective user training against today's cybercrimes. They can replicate thousands of scenarios for beginners and professionals alike, and offer a great opportunity in the academic sector to test theoretical information in practice.

One of the main challenges in building a cyber range is the complexity of the process. A cyber range typically involves a simulated network environment with various virtual machines, servers, and other components that are designed to emulate real-world cybersecurity scenarios. Designing and building such an environment can be a time-consuming, expensive, and technically challenging task, requiring expertise in areas such as networking, virtualization, development, and security. Additionally, maintaining and updating a cyber range can be an ongoing challenge, as new cyber threats and vulnerabilities emerge and need to be incorporated into the training scenarios. As a result, many organizations may struggle to build and maintain their cyber ranges, leading to a lack of effective cybersecurity training and educational resources [3].

A particular challenge in building a cyber range platform is providing an effective interface for administrators to control the whole platform and on the other side guide users through the prepared scenarios. This paper addresses these challenges and introduces the reader to the Brno University of Technology Cyber Arena (BUTCA) administration interface. The main function of this application is to provide administrators with an easy-to-use environment for controlling scenario creation and restoration to their original state, user management, and basic connectivity to the OpenStack virtualization platform through instances that are created when the user launches the selected scenario. Additionally, the application allows administrators to monitor and track the total cloud resources available at BUTCA. All of these functions are managed from a central application without the need to connect to other tools.

II. Related Work

Currently, several analyses are available that compare the various properties and functions of different types of CRs. In literature, at least 6 quality publications have been found that address this topic in the time span between 2013 and 2021. Most of the publications are focused on general CR types [4], [5]. Papers by Holm [6] and Kucek [7] are particularly focused on CRs with the ICS (Industrial Control System) and CTF (Capture the Flag) support. The most detailed article by Chouliaras [8] presents a review of CRs and, in addition to an analysis of the available literature, includes structured interviews with agencies and academic institutions that directly own and develop their CR. Chouliaras's article generally provides a comprehensive analysis of approximately 25 CRs and the description of 10 platforms is further expanded to include supported features and tools used. Although the number of CRs examined is the smallest compared to other research, the treatment is the most comprehensive of all. On the other hand, Yamin's research [1] examined the largest number of CRs (about 100), but the analysis only compared the basic features of the platforms.

The features relevant to the BUTCA platform presented in this paper are the custom administration software and the type of environment built by the CR. Table I provides a summary of basic information about CR platforms, including BUTCA (last row). The overview is based on [2] and has been extended by [8]–[28]. A total of 8 CRs were excluded from the overview due to the unavailability of information.

DOI 10.13164/eeict.2023.121

TABLE I: Overview of CR platforms basic information

Cyber range operator	Administration SW	Environment
KYPO CRP	✓	EM
Florida Cyber Range	✗	–
Virginia Cyber Range	✓	EM
Cylab.be	✓	S
AIT Cyber Range	✓	HCP
Norwegian Cyber Range	✓	EM, HCP, S
JYVSECTEC	✓	EM, HCP, S
CRATE	✓	HCP
Silensec Cyber Range	✓	–
CYBERIUM (Fujitsu)	✓	EM
DECIDE (NUARI)	✓	–
Hack the Box	✓	EM
TryHackMe	✓	EM
Georgia Cyber Range	–	EM
Cyberbit Cyber Range	✓	S
Airbus Cyber Range	✓	S
IBM X-Force C-TOC	✓	–
PwC Cyber Arena	✓	–
Cybexer	✓	EM
Raytheon Cyber Range	✓	S
CR14 NATO Cyber Range	✓	EM, HCP, S
RangeForce	✓	EM
BUTCA	✓	EM, HCP, I

Legend:
EM: Emulation HCP: Hybrid/Cyberphysical S: Simulation
I: Industry –: Information not available

The comparison above shows that most cyber ranges use proprietary software for administration and their environment is emulation-based. We have expanded our focus to include HCP and are the only ones listed to cover the industrial sector.

III. Proposed Solution

The proposed platform was created as part of an applied research project (VI20192022132), whose main goal was to develop a platform for research, testing, and education in cybersecurity. In terms of education, our primary focus was to support teaching for our Information Security study program, but we also extended our scope to high schools.

Users primarily log in to the application using a Microsoft account. There is also an option to register a new account using any email address, such as for invited guests who do not have a university account. All users are further divided into groups, typically by study courses or organization, to determine which game scenarios are available to them.

All scenarios have a maximum number of players and can also be limited by time. If an administrator starts a scenario (it is disabled by default), a user can only join if there is an available slot and can only run one scenario instance until they finish or run out of time. Players complete individual tasks and take a final test to verify their knowledge. Once the test is completed, the player's current ranking is displayed along with their total score. Players can also view a scoreboard and table of other players in the scenario during and after the game. Additionally, the administrator can remove a user instance at any time, view a log of incorrect answers for a given scenario, or view the scoreboard of all users' scores.

A. High-Level Architecture

The BUTCA platform consists of 3 components: application server, database, and Ansible AWX[1]. All of these components run in a virtual machine (VM) on top of the main part, which is the OpenStack cloud computing platform. The deployment and communication flow of all components are illustrated in Figure 1. Communication between the client's web browser and the application server is based on the Model-View-Controller (MVC) architectural pattern.

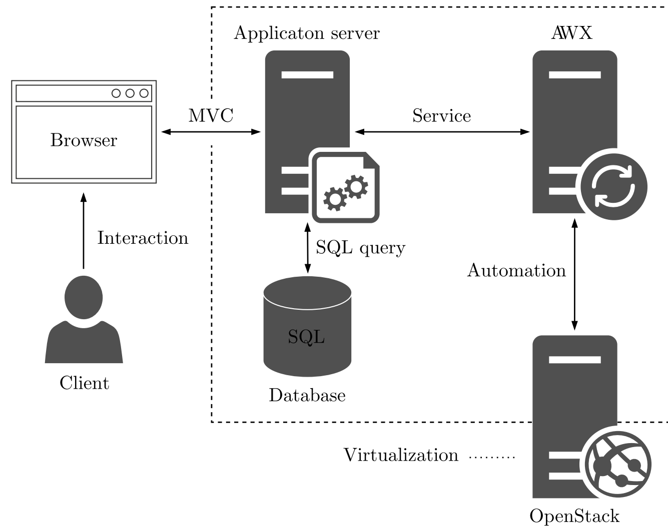


Fig. 1: High-level diagram of key components

B. Implemented Platform

We host all components of the BUTCA platform ourselves, including the web application for users and administrators, logging and monitoring, and the educational scenarios that we have created. One of the key advantages of our solution is the easy creation of scenarios and their gamification (story with prologue and epilogue, competitions, etc.). Our goal was to simplify the entire process for all users, so we developed a unified interface that allows administrators to manage all scenarios from a single application. Each game instance for connected users is composed of several tasks in a predefined linear scenario. Each task has a specific assignment that is either educational (more instructional) or part of the gamification story (more challenging). Tasks contain hints, the use of which may be penalized by the loss of points for the given task, and the last hint reveals the correct answer, but with a loss of 100% points. Furthermore, additional resources can be attached to the task, such as PDF files, multimedia files, ZIP archives, etc. The primary objective of each task is to find the flag, which represents the correct answer, similar to classic CTF (Capture the Flag) games. The user interface components described above are shown in Figure 2.

[1]The automation of complex processes is handled by the Virtualization Stack project in Ansible AWX, which is an open-source automation task engine with API (Application Programming Interface) developed by Red Hat.

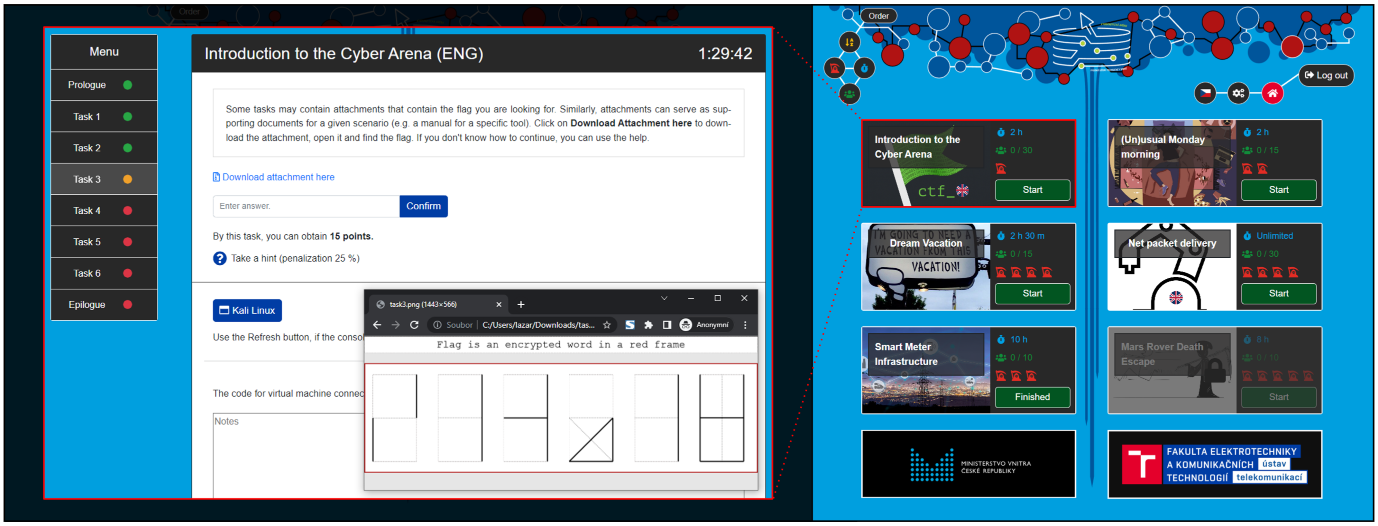


Fig. 2: User interface of the BUTCA (left – scenario instances, right – main dashboard)

IV. Testing and Evaluation

In total, the following 6 playtests with students were organized from Dec. 2021 to Dec. 2022:

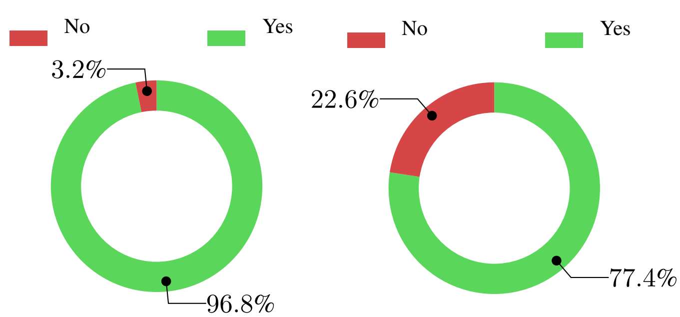
1) 47 students, FEEC Brno University of Technology.
2) 10 students, Secondary Technical School in Třebíč.
3) 28 students, Brno Grammar School, tř. Kpt. Jaroše.
4) 60 students, FEEC Brno University of Technology.
5) 42 students, Secondary Technical School in Třebíč.
6) 29 students, Tampere University (TAU).

During all playtests, logging and monitoring were conducted to track the following: incorrect answers, hints used, time taken to complete each task and scenario, final test points, overall scores of students, and feedback provided.

Based on the testing results, the difficulty was appropriately set and did not require any modifications in the future. Most of the students completed the tasks with at least one hint. However, one task needed a correction in the correct answer, which was a Base64 decoded string. According to the recorded incorrect answers, most of the students entered the correct result but in the wrong format, such as using lowercase instead of uppercase or subtracting spaces. The reason for these errors was that the students used different decoding tools, and some of them decoded the string differently than its original version. Therefore, the encoded string was modified to a single word consisting of only numbers and a special character to avoid such variations when using different tools.

Furthermore, we have established research cooperation with Tampere University to collect feedback based on the Attention, Relevance, Confidence, Satisfaction (ARCS) model [29]. For a better perspective of the collected feedback, we use the Reduced Instructional Materials Motivation Survey (RIMMS) version of the Keller's original survey, validated by [30], in which we have modified the survey items to fit our context. In the post-survey, students also assess how various CTF properties and CTF tasks affected their satisfaction, and how meaningful these properties and tasks were for their learning. In addition, they also assess how their general interest in cybersecurity changed due to these tasks and properties, which provides valuable feedback for teachers.

Surveys conducted during the playtesting with educational institutions investigate the interest in the BUTCA platform and cybersecurity in general, as well as the difficulty and quality of each task customized for particular scenarios. As an example, one result from a survey completed by 31 students from Secondary Technical School in Třebíč is shown in Figure 3. Charts in the figure indicate that students would welcome more playing sessions at the BUTCA, and a higher number of students would be satisfied if the playing sessions could take place on their campus without the need to commute to the Brno University of Technology.



(a) (b)

Fig. 3: Would you be interested in taking part in further testing of the BUTCA at (**a**) Secondary Technical School in Třebíč / (**b**) Brno University of Technology in the future?

In the future, we plan to include Czech students in research with Tampere University that focuses more on evaluating the effectiveness of cybersecurity education using cyber ranges in comparison to conventional teaching and learning, such as common lab tasks using VirtualBox or VMware.

V. Results and Discussion

From a technical point of view, our solution has passed all the tests successfully. However, there were some issues related to the user interface (UI) and user experience (UX) based on feedback from the involved students. Specifically, these include the organization of the help section, multi-level user access to the platform, and more detailed description of scenarios in the main menu. We will take all of these suggestions into account in the future development and focus on improving the UI and UX of the BUTCA.

Although the originally proposed solution was focused on supporting cybersecurity teaching at FEEC BUT, the idea of expanding educational scenarios to high schools arose during the platform development. Additionally, the user interface of BUTCA has been translated into English, opening up potential international collaborations, such as the current collaboration with the Tampere University. With increasing interest from other universities and high schools, we plan to redesign the current architecture of BUTCA to meet the CRaaS (Cyber Range as a Service) model, enabling students to play CTF scenarios directly from their university or high school. Our multi-level approach to cybersecurity education does not require any significant technical changes, and its implementation involves only creating scenarios with appropriate difficulty levels for the targeted level of study.

Based on the results achieved, new questions and suggestions for improving the current state of our solution have arisen. Future research and development will require further testing. By expanding the target group to include high schools, we can carry out testing with more subjects. In addition, establishing long-term cooperation with the Tampere University can help to continuously improve BUTCA, obtain more feedback, and collect research data.

Acknowledgment

We would like to express our gratitude to the Tampere University for their cooperation in organizing the playtesting sessions of the educational scenarios and for given feedback, namely to Tiina Schafeitel-Tähtinen and Marko Helenius from the NISEC research group.

References

[1] M. M. Yamin, B. Katt, and V. Gkioulos, "Cyber ranges and security testbeds: Scenarios, functions, tools and architecture," Computers & Security, vol. 88, p. 101636, Jan. 2020, doi: 10.1016/j.cose.2019.101636.

[2] I. Priyadarshini, "Features and Architecture of The Modern Cyber Range: A Qualitative Analysis and Survey," M.S. thesis, University of Delaware, Newark, DE, USA, 2018. [Online]. Available: http://udspace.udel.edu/handle/19716/23789/.

[3] T. Stodulka and R. Fujdiak, "Building the Cyber Range platform with cloud computing technology," Brno, Czechia: Brno University of Technology, 2022.

[4] J. Davis and S. Magrath, "A Survey of Cyber Ranges and Testbeds Executive," apps.dtic.mil. https://apps.dtic.mil/sti/citations/ADA594524 (accessed Feb. 24, 2023).

[5] E. Ukwandu et al., "A Review of Cyber-Ranges and Test-Beds: Current and Future Trends," Sensors, vol. 20, no. 24, p. 7148, Dec. 2020, doi: 10.3390/s20247148.

[6] H. Holm, M. Karresand, A. Vidström, and E. Westring, "A survey of industrial control system testbeds," presented at the 20th Nordic Conference, NordSec 2015, Stockholm, Sweden, Oct. 2015, pp. 11–26, doi: 10.1007/978-3-319-26502-5_2.

[7] S. Kucek, M. Leitner, "An empirical survey of functions and configurations of open-source capture the flag (ctf) environments," Journal of Network and Computer Applications, vol. 151, p. 102470, Feb. 2020, doi: 10.1016/j.jnca.2019.102470.

[8] N. Chouliaras, G. Kittes, I. Kantzavelou, L. Maglaras, G. Pantziou, and M. A. Ferrag, "Cyber Ranges and TestBeds for Education, Training, and Research," Applied Sciences, vol. 11, no. 4, p. 1809, Feb. 2021, doi: 10.3390/app11041809.

[9] Z. Eichler, R. Ošlejšek, and D. Toth. "KYPO: A Tool for Collaborative Study of Cyberattacks in Safe Cloud Environment," Human Aspects of Information Security, Privacy, and Trust, Cham, Switzerland: Springer, 2015, pp. 190–199, doi: 10.1007/978-3-319-20376-8_17.

[10] E. M. Raybourn, M. Kunz, D. Fritz, and V. Urias, "A zero-entry cyber range environment for future learning ecosystems," Cyber-Physical systems security, Cham, Switzerland: Springer, 2018, pp. 93–109, 2018, doi: 10.1007/978-3-319-98935-8_5.

[11] N. M. Radziwill, "Virginia cyber range," Software Quality Professional, vol. 19, no. 4, p. 46, Sep. 2017.

[12] M. Leitner, et al., "AIT cyber range: flexible cyber security environment for exercises, training and research," presented at the European Interdisciplinary Cybersecurity Conference, Nov. 2020, pp. 1–6, doi: 10.1145/3424954.3424959.

[13] M. Kianpour, S. Kowalski, E. Zoto, C. Frantz, and H. Øverby, "Designing serious games for cyber ranges: a socio-technical approach," presented at the 2019 IEEE European symposium on security and privacy workshops (EuroS&PW), Jun. 2019, pp. 85–93, doi: 10.1109/EuroSPW.2019.00016.

[14] J. Piispanen, "Technical specification for federation of cyber ranges," M.S. thesis, JAMK University of Applied Sciences, Jyväskylä, Finland, 2018. [Online]. Available: https://urn.fi/urn:nbn:fi:amk-2018121722010.

[15] T. Gustafsson and J. Almroth, "Cyber range automation overview with a case study of CRATE," presented at the 25th Nordic Conference, NordSec, Nov. 2020, pp. 192–209, doi: 10.1007/978-3-030-70852-8_12.

[16] "CYBER RANGES: Cyber Security Exercises for Training and Capability Development," cyberranges.com. https://www.cyberranges.com/ (accessed Feb. 25, 2023).

[17] K. Hara, "Cyber range CYBERIUM for training security meisters to deal with cyber attacks," Fujitsu Sci. Tech. J, vol. 55, no. 5, pp. 59–63.

[18] "DECIDE®: Cyber Security Live Exercises," nuari.org. https://nuari.org/decide (accessed Feb. 25, 2023).

[19] "Introduction to Hack The Box," help.hackthebox.com. https://help.hackthebox.com/en/articles/5185158-introduction-to-hack-the-box (accessed Feb. 25, 2023).

[20] "TryHackMe — Hacktivities," tryhackme.com. https://tryhackme.com (accessed Feb. 25, 2023).

[21] "Cyber Range – Georgia Cyber Innovation & Training Center," gacybercenter.org. https://www.gacybercenter.org/services/cyber-range/ (accessed Feb. 25, 2023).

[22] "Cyberbit – Cyber Range Platform leading provider," cyberbit.com. https://www.cyberbit.com/ (accessed Feb. 25, 2023).

[23] "CyberRange – Airbus Defence and Space Cyber," cyber.airbus.com. https://www.cyber.airbus.com/cyberrange/ (accessed Feb. 25, 2023).

[24] "IBM Security X-Force Cyber Range," ibm.com. https://www.ibm.com/services/security-operations-center (accessed Feb. 25, 2023).

[25] "Cyber Arena," cyberarena.pwc.com. https://cyberarena.pwc.com (accessed Feb. 25, 2023).

[26] "Cyber Range as a Service – CybExer Technologies," cybexer.com. https://cybexer.com/solutions/cyber-range-as-a-service/ (accessed Feb. 25, 2023).

[27] "CR14 Ranges," cr14.ee. https://www.cr14.ee/ (accessed Feb. 25, 2023).

[28] "Cyber Skills Training," rangeforce.com. https://www.rangeforce.com (accessed Feb. 25, 2023).

[29] J. M. Keller, "Motivational Design for Learning and Performance." Boston, MA: Springer US, 2010, doi: 10.1007/978-1-4419-1250-3.

[30] N. Loorbach et al., "Validation of the Instructional Materials Motivation Survey (IMMS) in a self-directed instructional setting aimed at working with technology: Validation of the IMMS," British Journal of Educational Technology, vol 46, no 1, pp. 204–218, 2015, dot: 10.1111/bjet.12138.

Network Toolkit for exploiting Internet Protocol version 6 Security Vulnerabilities

1st Viet Anh Phan
Department of Telecommunications, FEEC
Brno University of Technology
616 00 Brno, Czech Republic
243760@vut.cz

2nd Jan Jeřábek
Department of Telecommunications, FEEC
Brno University of Technology
616 00 Brno, Czech Republic
jan.jerabek@vut.cz

***Abstract*—This work has developed a network toolkit in Python to automatically carry out several network attacks and security vulnerability testing when operating IPv6. Specifically, the program can be used to launch specified practical types of attacks according to the user's direction such as DoS (Denial of Service), dumping servers, spoofing or bypassing the firewall. The essence of these attacks is based on the inherent vulnerabilities of Extension headers, and protocols such as ICMPv6 (Internet Control Message Protocol), DHCPv6 (Dynamic Host Configuration) Protocol, which are the most important protocols in the IPv6 network operation. In addition, the analysis and illustrations have been presented, which can help network analysts to have a clearer understanding of the potential dangers that could arise from implementation of IPv6. From there, they can propose appropriate solutions to eliminate or mitigate damage when attacks occur.**

***Index Terms*—IPv6, ICMPv6, Toolkit, Network security, Attack, Python, Kali Linux, Scapy**

I. INTRODUCTION

IPv4 recently reached its limit in terms of the number of addresses. There are also many limitations that cannot be completely overcome when working with protocols under IPv4. Therefore, IPv6 has been and is being widely deployed in the infrastructure facilities of many internet service providers (ISP) and also content delivery networks (CDN). As a matter of course, the novelty of IPv6 can pose difficulties for security systems as well as network analysts in detecting and preventing potential security vulnerabilities or attacks. The potential threats have prompted the development of tools that can simulate network attacks. From there, network experts can analyze, acquire experiences, and take countermeasures to protect the system from service disruption or loss of information data or unscheduled costs.

To run this toolkit, it is necessary to have Python 3.x and the latest version of Scapy library (obtained from GitHub link [1]). Scapy is a Python-based packet manipulation library that allows users to create, send, and capture network packets. It provides a simple and powerful interface to interact with different layers of network protocols, including Ethernet, IPv6, ICMPv6, DNS (Domain Name System), TCP (Transmission Control Protocol) and UDP (User Datagram Protocol). Kali Linux, which is a popular and powerful Linux-based operating system, is used for sending packets created by this Python network toolkit and capturing messages due to specific purposes.

There are several IPv6 toolkits available that are specifically designed to help with IPv6 network testing and analysis such as Thc-IPv6 [2], Chiron [3] and Ipv6-toolkit [4]. These tools have some limitations. For example, when performing flooding attacks, Thc-ipv6 only has a low packet sending rate per second. This leads to the fact that it does not really disable victims or make them unresponsive in many cases. Chiron only runs on the Python 2.7 platform, so it is not widely used due to lacking many language features, community support, and compatibility with modern systems today. IPv6-toolkit is probably too complicated in the way of inserting parameters, which requires users to thoroughly research each detail if they want to use it properly.

The newly designed network toolkit is planned to fix some weaknesses of above mentioned tools and provide some additional functionalities in particular cases. Specifically, this designed toolkit can launch DoS attack with the maximum rate 70000 packets/second (compared to 20000 packets/second with the same type and size in case of Thc-IPv6). Moreover, user can easily observe the possible responses and attacking results shown in the terminal when running the toolkit. This feature is not available at all other mentioned toolkits even though they are more complicated for user to launch. The toolkit is still being further developed and its current source code is available at https://github.com/vafekt/IPv6-toolkit.

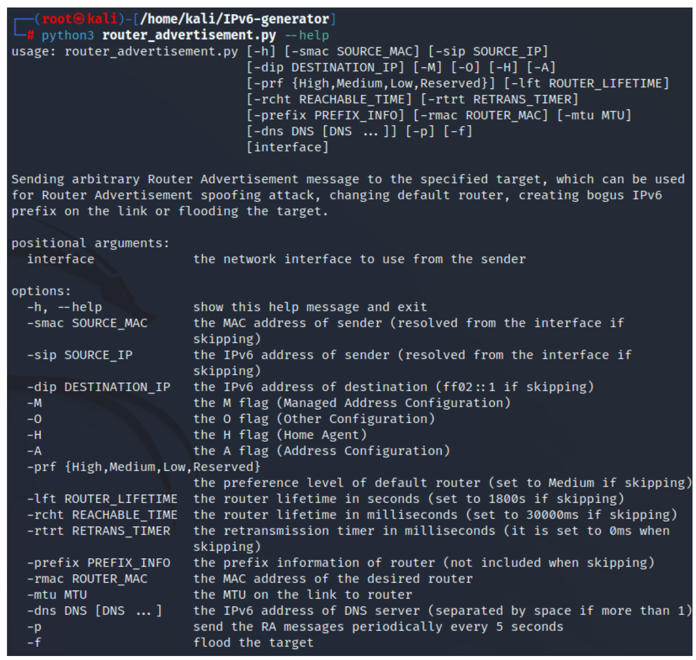
II. BASIC DESCRIPTION OF THE DEVELOPED NETWORK TOOLKIT

At present, this network toolkit is comprised of the following modules, which can be used for specific purposes such as sending, sniffing, detecting and attacking:

- **ping.py**: A tool that sends ICMPv6 Echo Request messages from specified source to destination. Users are able to define how many packets to send, the size of data, option to send the file and the Ping Flood attack.
- **routing_header.py**: A tool that sends ICMPv6 Echo Request messages from specified source through defined intermediate hops to the destination. The Flood attack is included to slow down every hosts on the way to the destination.

- **destination_header.py**: A tool that sends SYN or ICMPv6 packet with several options of Destination Option header such as 1x Destination Option header; 1x Destination Option header with hidden data; 3x Destination Option headers; 100x Destination Option headers and 200x Destination Option headers for checking abilities to bypass the firewall.
- **fragment_header.py**: A tool that sends SYN or ICMPv6 packet with several options of Fragment header including 1x atomic fragment (with same id and different id); 3x atomic fragment (with same id and different id); 100x atomic fragment (with same id and different id); tiny fragments and overlapping fragments. It is used for checking abilities to bypass the firewall.
- **redirect.py**: A tool that sends Redirect message to a victim host, causing it to update its routing table with a new, incorrect route to a destination host. The flooding option is also included to cancel the legitimate routing permanently.
- **mldv2_query.py**: A tool that sends several types of Multicast Listener Discovery version 2 (MLDv2) Query message to discover multicast listeners in the network or pretend to be the multicast router. It also has the option of flooding attack to make the target unresponsive.
- **mldv2_report.py**: A tool that sends Multicast Listener Discovery version 2 (MLDv2) Report message to inform routers about the multicast groups that have active listeners on a network segment, add or delete multicast listeners. DoS is included as option flooding attack.
- **covert_channel.py**: A tool that performs hidden communication between two specified entities by using Covert Channel in Extension headers. It has option to encrypt the file with AES (Advanced Encryption Standard) or DES (Data Encryption Standard).
- **detect_alive.py**: A tool that detects active hosts on the attached local link, including their IPv6 addresses and link-layer addresses.
- **detect_new.py**: A tool that detects new hosts joining the attached local link based on Duplicate Address Detection (DAD) process.
- **neighbor_solicitation.py**: A tool that sends Neighbor Solicitation message to specified target for resolving link-layer address and status of that victim, together with option to flood it.
- **neighbor_advertisement.py**: A tool that answers every Neighbor Solicitation message with a falsified Neighbor Advertisement message for spoofing address resolution.
- **router_solicitation.py**: A tool that sends arbitrary Router Solicitation message to specified target, with option to flood.
- **router_advertisement.py**: A tool that sends arbitrary Router Advertisement message to specified target with an aim to spoof attack, changing the default router, creating bogus IPv6 prefix on the link or poisoning routing entries. The option to flood is included to cause DoS on the target.
- **prevent_autoconfiguration.py**: A tool that prevents every host on the local link from auto-configurating its global IPv6 addresses. Therefore, these hosts cannot communicate with the external network.
- **implant_mtu.py**: A tool that implant the Maximum Transmission Unit (MTU) to a specified target, so the victim can only transfer data up to this defined MTU.
- **smurf.py**: A tool that triggers smurf attack to flood a specified target.
- **flood_dhcpv6.py**: A tool that kills all DHCPv6 servers with enormous falsified DHCPv6 Solicit messages and option Rapid Commit.
- **dhcpv6_server.py**: A tool that operates as a fake DHCPv6 server.

To launch the program properly, users must access to the directory of the toolkit, choose one of the tools and run it as root. When starting the module, it is necessary to insert specific parameters (users must define at least the network interface to use). At any time when needing help, users can run the particular tool with **-help** switch for more information, which is shown in Fig. 1 for the case of tool **router_advertisement.py**.



```
┌──(root㉿kali)-[/home/kali/IPv6-generator]
└─# python3 router_advertisement.py --help
usage: router_advertisement.py [-h] [-smac SOURCE_MAC] [-sip SOURCE_IP]
                               [-dip DESTINATION_IP] [-M] [-O] [-H] [-A]
                               [-prf {High,Medium,Low,Reserved}] [-lft ROUTER_LIFETIME]
                               [-rcht REACHABLE_TIME] [-rtrt RETRANS_TIMER]
                               [-prefix PREFIX_INFO] [-rmac ROUTER_MAC] [-mtu MTU]
                               [-dns DNS [DNS ...]] [-p] [-f]
                               [interface]

Sending arbitrary Router Advertisement message to the specified target, which can be used
for Router Advertisement spoofing attack, changing default router, creating bogus IPv6
prefix on the link or flooding the target.

positional arguments:
  interface             the network interface to use from the sender

options:
  -h, --help            show this help message and exit
  -smac SOURCE_MAC      the MAC address of sender (resolved from the interface if
                        skipping)
  -sip SOURCE_IP        the IPv6 address of sender (resolved from the interface if
                        skipping)
  -dip DESTINATION_IP   the IPv6 address of destination (ff02::1 if skipping)
  -M                    the M flag (Managed Address Configuration)
  -O                    the O flag (Other Configuration)
  -H                    the H flag (Home Agent)
  -A                    the A flag (Address Configuration)
  -prf {High,Medium,Low,Reserved}
                        the preference level of default router (set to Medium if skipping)
  -lft ROUTER_LIFETIME  the router lifetime in seconds (set to 1800s if skipping)
  -rcht REACHABLE_TIME  the router lifetime in milliseconds (set to 30000ms if skipping)
  -rtrt RETRANS_TIMER   the retransmission timer in milliseconds (it is set to 0ms when
                        skipping)
  -prefix PREFIX_INFO   the prefix information of router (not included when skipping)
  -rmac ROUTER_MAC      the MAC address of the desired router
  -mtu MTU              the MTU on the link to router
  -dns DNS [DNS ...]    the IPv6 address of DNS server (separated by space if more than 1)
  -p                    send the RA messages periodically every 5 seconds
  -f                    flood the target
```

Fig. 1: Manual page of router_advertisement.py tool.

III. Brief Description of the Testing Environment

The testing is implemented in the virtual machine environment powered by GNS3. There are overall 7 devices including 2 Cisco routers; 1 Ethernet switch; 1 Kali Linux PC with version 2022.4; 1 Windows 10 Home PC and 2 Ubuntu PCs with version 22.04.1 LTS (Long-term Support). The network topology is depicted below in Fig. 2.

Router R1 and R2 are the legitimate default routers on their network segments (`2001:db8:abcd:1::/64` and `2001:db8:abcd:3::/64`, respectively). These two

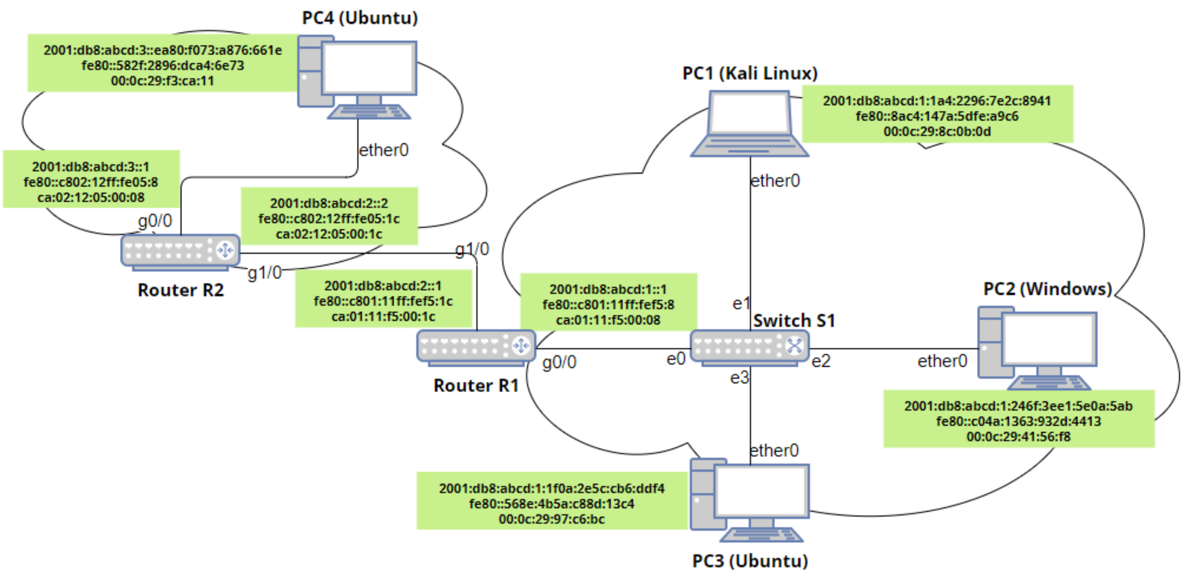


Fig. 2: The testing network topology with address specification (/64 prefix used in all subnets).

routers provide Router Advertisement to clients for address auto-configuration. The PCs from number 1 to number 3 reside on the local link of router R1, while PC4 locates on the network of router R2. All PCs are the clients in the model, and PC1 (Kali Linux) takes on the job of being an attacker in specific situations, which will be mentioned later.

In the scope of the article, two tools including **detect_alive.py** (section IV) and **neighbor_advertisement.py** (section V) from the toolkit are presented briefly to see how they impact victims and/or the network operations.

IV. Brief Description of the detect_alive.py Tool

Before executing any type of attack, it is essential to get information about active hosts on the attached link and about the target. Ping sweep or similar brute force scans is usually applied in IPv4, but conducting this type of scan of all IPv6 addresses is impractical due to the vast address space of IPv6 subnets. Instead, this stage can be carried out by transmitting ICMPv6 Echo Request messages (or similar packets) to the all-nodes multicast address.

However, not all nodes respond to the normal ICMPv6 Echo Request (notably, machines running Windows system often ignore this type of packet). Therefore, the idea of sending an additional type of packet, which is called malformed ICMPv6 Echo Request, is implemented. In this packet, an unknown option of Destination header is inserted (depicted in the yellow frame of Fig. 3) to make hosts unable to process the packet. Then, all hosts should send a Parameter Problem response (shown as the red frame of Fig. 3) back to the attacker. This has been defined in specification RFC 4884 [5]. As a result, information such as IPv6 address and MAC address of victims are exposed.

That is what the tool **detect_alive.py** performs. In the scenario (Fig. 2), it is taken into account that all devices are in active mode. From PC1 (Kali Linux), which is the attacker,

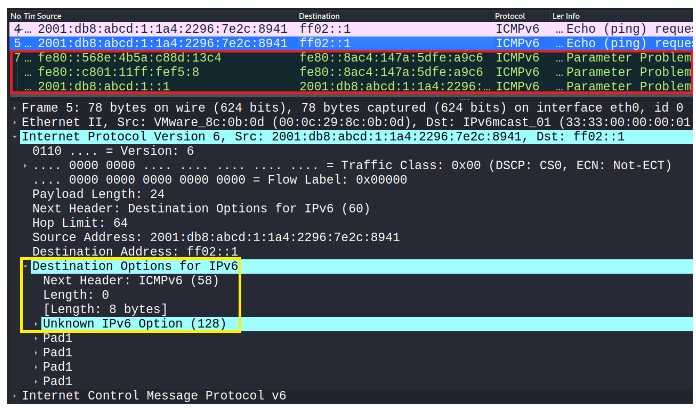


Fig. 3: Illustration of captured malformed ICMPv6 Echo Request packet and its responses.

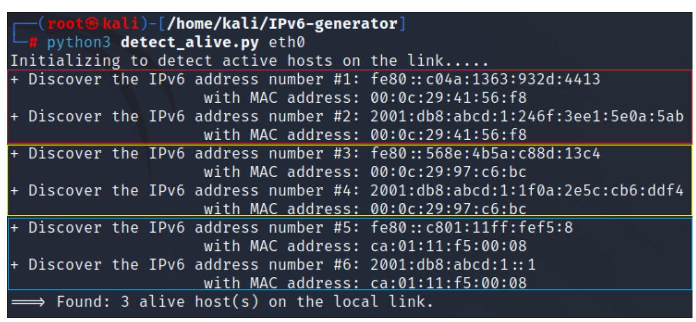


Fig. 4: The operation and result of detect_alive.py tool after running.

the tool is launched as shown in the Fig. 4. The toolkit sends ICMPv6 Echo Request packets using both IPv6 link-local and global source addresses in order to get the link-local and global ones from victims.

After launching the attack, there are 3 hosts found on the local link (Fig. 4). In the local network containing attacker's PC1, there are 3 hosts including Router R1, PC2 (Windows), and PC3 (Ubuntu). The red frame contains address information about PC2, the yellow frame is from PC3, and the blue frame is from router R1. Therefore, the program has operated correctly. Obtained addressing information could be used when using other tools.

V. Brief Description of the neighbor_advertisement.py Tool

This tool executes Neighbor Advertisement spoofing attack. Specifically, if the attacker (PC1) responds to a Neighbor Solicitation (NS) message with a falsified Neighbor Advertisement (NA) that contains his own link layer address, he can disrupt the process of resolving link layer addresses.

Since the victim (router R1, PC2 or PC3) accepts this NA packet, it will start sending all data link frames to the MAC address of the attacker. If the attacker goes a step further and uses a falsified NA message to spoof the destination node, he can carry out a man-in-the-middle attack, which allows him to intercept and read all data transmitted between the two parties involved in the conversation, as depicted in Fig. 5. The attack in progress is shown in Fig. 6.

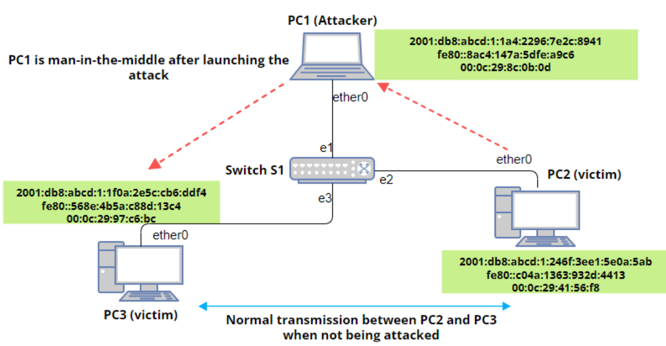


Fig. 5: Illustration of Neighbor Advertisement spoofing attack in the scenario.

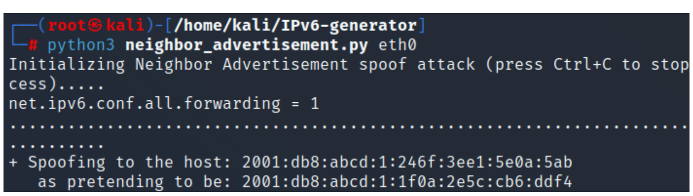


Fig. 6: The running program of detect_alive.py tool.

In this scenario, when first communicating with PC3, PC2 sends Neighbor Solicitation message for achieving the link layer address of PC3. The attacker PC1 spoofs the PC2, as the result is described in Fig. 6. The attacker then pretends to be PC3. By allowing traffic forwarding (`sysctl -w net.ipv6.conf.all.forwarding=1`) in Kali operating system, the attacker now becomes man-in-the-middle, which means PC2 cannot communicate directly with PC3 but the PC1 (attacker) will receive the data from PC2 and forward it to PC3.

Besides, it is worth noting that this attack can occur in two directions, meaning that PC1 (the attacker) spoofs PC3 in the opposite direction and pretends to be PC2. As a result, the entire information exchange process between the two PCs is completely interfered by the attacker. Similarly, we can use this attack to pretend to be legitimate default gateway of the network and become man-in-the-middle for whole network.

As can be seen from the Listing 1, the neighbor cache of PC2 (Windows) are observed to change as soon as PC2 and PC3 start sending messages to each other. The same thing happens to the neighbor caches of PC3. After the attack happens, the link layer address of PC3 changes from the real one (`00:0c:29:97:c6:bc`) to a new address (`00:0c:29:8c:0b:0d`), which is actually the link layer address of the attacker PC1. This allows the attacker to redirect all IPv6 connections through his machine.

```
# Before the attack
C:\WINDOWS\system32>netsh interface ipv6 show neighbor

Internet Address                              Physical Address
-------------------------------------------   -----------------
2001:db8:abcd:1::1                            ca:01:11:f5:00:08
2001:db8:abcd:1:1a4:2296:7e2c:8941            00:0c:29:8c:0b:0d
2001:db8:abcd:1:1f0a:2e5c:cb6:ddf4            00:0c:29:97:c6:bc

# After the attack
C:\WINDOWS\system32>netsh interface ipv6 show neighbor

Internet Address                              Physical Address
-------------------------------------------   -----------------
2001:db8:abcd:1::1                            ca:01:11:f5:00:08
2001:db8:abcd:1:1a4:2296:7e2c:8941            00:0c:29:8c:0b:0d
2001:db8:abcd:1:1f0a:2e5c:cb6:ddf4            00:0c:29:8c:0b:0d
```

Listing 1: The neighbor cache in PC2 (Windows) before and after the happening attack (outputs shortened).

Conclusions

The structure of the network toolkit for exploiting IPv6 security vulnerabilities has been briefly described with concrete modules. The toolkit is still under development, yet, it has removed some drawbacks of other available attacking toolkits. Two modules (section IV and V) have been briefly presented to prove the workability of the toolkit. In the future, this toolkit is continuing to be developed with more features (types of attack, related options) and to be more user friendly.

References

[1] P. Biondi, Scapy, [online] Available: http://www.secdev.org/projects/scapy.

[2] M. van Hauser. THC IPv6 Attack Toolkit, [online] Available: http://www.thc.org/thc-ipv6/

[3] Chiron, [online] Available: https://www.secfu.net/tools-scripts/

[4] IPv6-toolkit, [online] Available: https://www.kali.org/tools/ipv6-toolkit/

[5] Extended ICMP to Support Multi-Part Messages, [online] Available: https://www.rfc-editor.org/rfc/rfc4884

[6] J. Jeřábek, *Pokročilé komunikační techniky*. Skriptum FEKT Vysoké učení technické v Brně 2022, s.1-174.

[7] Kurose, J. F., Ross, K. W., Computer networking: a top-down approach. 7th global ed. Essex: Pearson, 2017, 852 s. ISBN 978-1-292-15359-9.

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.

Blood pressure estimation using smartphone

1st Jan Šíma
Department of Biomedical Engineering
FEEC Brno University of Technology
Brno, Czech Republic
E-mail: xsimaj02@vutbr.cz

2nd Andrea Němcová
Department of Biomedical Engineering
FEEC Brno University of Technology
Brno, Czech Republic
E-mail: nemcovaa@vut.cz

***Abstract*—This paper presents an experimental cuff-less measurement of systolic (SBP) and diastolic blood pressure (DBP) using smartphone. A photoplethysmographic signal (PPG) measured by a smartphone camera is used to estimate blood pressure (BP). This paper contains comparison of several machine learning (ML) methods for BP estimation. Filtering the PPG signal with a band-pass filter (0.5-12 Hz) followed by feature extraction and using Random Forest (RF) methods separately or as a weak regressor in adaptive boosting (AdaBoost) or bootstrap aggregating (Boosting) reached the best results according to Association for the Advancement of Medical Instrumentation (AAMI) and British Hypertension Society (BHS) standards among all regression ML models. The mean absolute error (MAE) and standard deviation (SD) of Bagging model were 4.532±3.760 mmHg for SBP and 2.738±3.032 mmHg for DBP (AAMI). This result meets the criteria of the AAMI standard.**

***Index Terms*—blood pressure estimation, cuff-less measurement of blood pressure, machine learning**

I. Introduction

Cardiovascular diseases (CDC) are currently the leading cause of death in the world. For example, in 2019, 17.9 million people died from CDC (32 % of all global deaths) [1]. The main cause is neglected prevention and early detection of symptoms leading to CDC. Some of the symptoms of CDC cannot be revealed at a routine check-up at the doctor's office. Therefore, there is a need for continuous long-term monitoring. Common and important predictors of CDC include hypertension. Its symptoms are high heart rate (HR), severe palpitations, fatigue, shortness of breath, etc. The diagnosis of hypertension is confirmed after a sustained blood pressure (BP) above 140/90 mmHg [2].

Therefore, the solution to this global problem is to increase CDC prevention globally by using smart wearable devices that can sense and analyze a given biological signals continuously.

In recent years, research and development in the field of BP measurement have focused on sensing and analyzing biological signals using smartphones and smartwatches. Their biggest advantage is that they are user-friendly and affordable. Plus, they are widespread and comfortable, because the BP measurement is cuff-less.

There were 6.26 billion smartphone users in 2021, the majority of the world's population [3]. If every smartphone user could monitor their BP it would significantly improve CDC prevention or at least early diagnosis.

This paper describes an experimental measurement of BP using a smartphone and a comparison of the application of machine learning (ML) methods for estimating systolic blood pressure (SBP) and diastolic blood pressure (DBP). The biological signals that can be sensed by a smartphone to estimate BP is photoplethysmographic signal (PPG) and phonocardiographic signal (PCG).

Smartphones or watches are exactly the devices that are suitable for non-invasive, cuffless pressure monitoring. Recently, PPG signals have been used to estimate BP. When a LED-illuminated finger is videorecorded, the change in tissue volume is captured in the resulting video frames by varying the brightness of the image. Current BP influences the strength of pulse wave and its morfology. Methods for estimating BP from PPG signals can be divided into mathematical approaches and ML approaches. In this paper, only ML approaches are used.

The chapter II describes the entire experimental measurement procedure and analysis. As an experimental data for cuffless BP estimation, PPG and PCG signals were sensed using smartphone in-built sensors. Reference BP was sensed with a certified medical device. Furthermore, preprocessing, datasets description, signal quality estimation, feature extraction, and ML methods that were used to estimate SBP and DBP are presented in this chapter. In chapter III, the results of BP estimation using different ML approaches are compared. Error assessment of the regression models was performed based on standard of Association for the advancement of medical instrumentation (AAMI) and standard of the British hypertension society (BHS).

II. Data and methods

A. Experimental measurement

The aim of the experimental measurements was to obtain quality data from the mobile phone (PPG and PCG) and the reference device (BP).

A Samsung Galaxy S7 smartphone was chosen to capture the video recording (PPG) with simultaneous audio recording (PCG). The camera of the smartphone was set to full high definition (FHD) resolution (1920×1080 px) and sampling rate 60 fps. When measuring, the finger of the right hand is placed on the camera lens on the back of the phone. The LED is lit during the recording for better data quality.

The whole experimental measurement was carried out at the Department of Physiology, Faculty of Medicine, Masaryk University (MUNI). As a reference for BP measurement, a medically certified device Finapres NOVA was used. This device sensed the reference BP by the Penaz method from the middle finger of the left hand. In addition, electrocardiogram (ECG), respiration, HR and pulse oximetry signals were also measured with the Finapres NOVA.

For further use/research, a recording from an Empatica smart bracelet was taken on the right wrist. Empatica allows to record accelerometer data (ACC), PPG, HR, patient temperature and electrodermal activity (EDA).

Figure 1 illustrates the placement of all sensors on the body during experimental measurement.

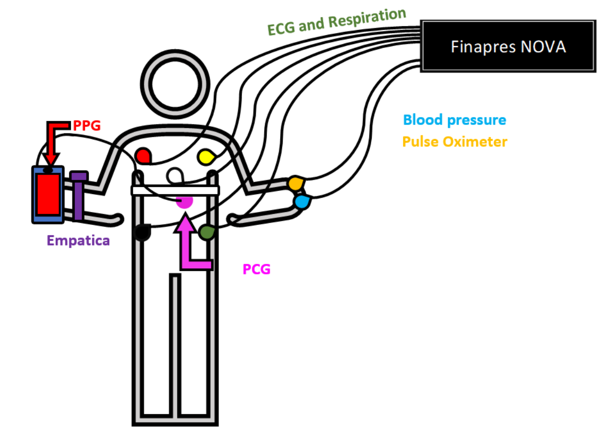


Fig. 1. Illustration of the placement of all sensors on the body during experimental measurement.

All measured subjects gave their informed consent. The experimental measurements were approved by the Ethical Committee Faculty of Electrical Engineering and Communication Technology (FEEC) For Biomedical Research Brno University of Technology (BUT).

The experimental measurement protocol is shown in Figure 2.

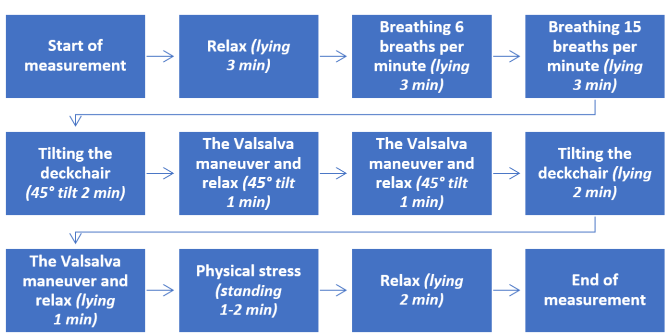


Fig. 2. The experimental measurement protocol.

During the measurements, various tasks were included to invoke changes in BP. The length of the protocol was set to 20–22 min as a compromise between comfort (e.g. finger is heated by the LED) and achievable data quality on the one hand and sufficient variability of BP data on the other hand. Significant variability in BP values is required for reliable training of ML model.

B. Preprocessing

For this study, it was necessary to have synchronized PPG signal from smartphone and reference BP values from Finapres NOVA. Nevertheless, the synchronization was done using PCG signal from smartphone and ECG signal from Finapres NOVA. Accurate synchronization is possible due to the temporal alignment of the electrical activation of the ventricles (R peak in ECG) and the mechanical activation of the ventricles (S1 heart sound in PCG).

The PPG signal was filtered by a 2nd order band-pass Butterworth filter with two different cutoff frequencies. The first dataset (D1) with cutoff frequencies of 0.5 Hz and 8 Hz [4]. The second dataset (D2) with cutoff frequencies of 0.5 Hz and 12 Hz. Everything below 0.5 Hz can be attributed to baseline wander, while everything above 12 Hz is high frequency noise [4].

Loading and synchronization were implemented in the Matlab programming environment.

C. Pulse wave detection

To detect individual PPG waves, almost all frequency components need to be suppressed except for the region around the 1st harmonic component (HR). This component is approximately between 0.5 Hz and 3 Hz. The PPG was filtered by finite impulse response (FIR) bandpass filter which cutoff frequencies were set to 0.5 Hz and 3 Hz and the length of impulse response to 2311 samples.

Individual PPG peaks were not detected, but the minima between them (hereafter referred to as ”onsets”) were detected. This was achieved in the Matlab environment by detecting local minima which minimum distance was set to 0.4 s.

D. Pulse wave quality assessment

To determine the quality of individual PPG waves, 5 conditions were used. A quality pulse wave was included in the dataset if:

- The calculated HR value was between 30 bpm and 180 bpm.
- The peak of the pulse was between 0.075 s and 0.35 s from the start of the pulse wave.
- Peak width was in 90 % of PPG wave height was less than 0.3 s.
- The start or end point after normalization was less than 40 % of the amplitude of the PPG wave.
- The correlation coefficient [%] between the PPG wave and the reference pulse wave was greater than a specified threshold. (90 % or 99 %, specified in Table I for both datasets).

E. Dataset

A total of 13 subjects, 8 males and 5 females, participated in the experimental measurement. Before each measurement, a questionnaire was completed with the volunteer. Based on this questionnaire, basic features such as age, weight, height and some health features were extracted.

The average age of measured volunteers was 39 years, the youngest participant was 20 years old and the oldest was 74 years old. The mean height was 173 cm with a standard deviation of 9 cm, with the shortest participant of 160 cm and the tallest participant measuring 194 cm. The mean weight was 82.7 kg with a standard deviation of 23.5 kg. The woman with the lowest height of 160 cm also had the lowest weight of 52 kg and the tallest man of 194 cm weighed 127 kg.

Among the subjects, there were three volunteers who were treated for high BP and took medication to control it. We were even able to capture data from a volunteer who was currently treated for coronary artery disease (male, 74 years old) and a volunteer who had a cardiac arrhythmia, specifically ventricular extrasystole (female, 70 years old).

The measured dataset contains about 15,000 PPG pulses, which were subsequently processed and their quality was determined. Based on the quality features, only those pulses that met certain criteria were selected for feature extraction.

Statistical summary of SBP and DBP for both dataset is shown in Table I. For clarification, e.g., the dataset name D2_99 refers to the dataset with band-pass (0.5-12 Hz) filtered signals with quality threshold of 99 %.

TABLE I
DATASETS STATISTICAL SUMMARY OF SBP AND DBP.

Datasets	BP	Mean[mmHg]	SD[mmHg]	Range[mmHg]
D1_90 (N = 8084)	SBP DBP	122.63 67.96	20.54 17.33	164.68 165.48
D2_99 (N = 1107)	SBP DBP	113.45 66.73	12.04 9.05	135.61 70.09

F. Feature extraction

The measurement involved the collection of health data in the form of a questionnaire. These data were also used as features for training the regression model [2] [5].

To properly train ML models, it is useful to have a large enough parameter space. The features should not correlate with each other and should have high variability.

In this paper, PPG from a smartphone was used to estimate BP. In the time domain, in each beat a few features were extracted (described below). Moreover, HR was extracted from each two adjacent PPG peaks [5].

Due to artifacts in PPG signals, the pulse amplitude can vary significantly and therefore cannot be used as a feature for BP estimation.

Each pulse was first normalized using max-min normalization. It is then divided into an anacrotic and a catacrotic part, from which the temporal features were extracted.

The extraction of time domain features includes intervals from the filtered pulse wave, its first derivative and second derivative. Thus, a total of 66 features including HR were extracted from the time domain [2] [6].

Figure 3 shows an example of a measured pulse wave PPG, where S/D denotes the systolic and diastolic regions and the number represents the percentage of the peak height.

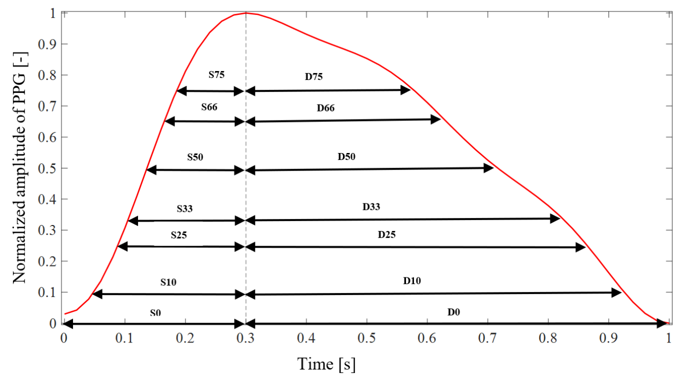


Fig. 3. PPG pulse wave with extracted features (S/D - systolic and diastolic area, number - percentage of the peak height).

10 qualitative features were also added to the feature field. These are the aforementioned correlation coefficients for the reference median pulse wave and average pulse wave, mean square error (MSE) between the reference pulse waves and the extracted PPG waves, entropy, etc. The total number of features was 95.

G. Regression approaches

The Python programming environment was used to implement the ML methods. Specifically, the PyTorch library was used.

Here is a list of ML regression approaches. Information about model parameters is added where appropriate.

Linear regression model (LRM), k–nearest neighbors (k–NN) algorithm with 5 nearest neighbors, decision tree (DT) with maximum depth 40 and mean absolute error (MAE) criterion, random forest (RF) with maximum depth 40, support vector regression (SVR) model with radial basis function (RBF) as kernel type of degree 3, ridge regularization (LRMR) with alpha parameter 1, adaptive boosting (AdaBoost) with weaker RF regressors (N=88), bootstrap aggregation (Bagging) with weaker RF regressors (N=100) and artificial neural network (ANN) with 1 hidden layer (100 perceptrons).

The special solution was a convolution neural network (CNN) with 2 convolutional layers (50 filters), 2 max pooling blocks, batch normalization and 2 fully connected (FC) layers (4520 and 2500 perceptrons). The input was the normalized PPG signal, its 1st derivative and 2nd derivative. These were concatenated in sequence. The FC layer received input from the health features (described above) and output from the convolutional layers.

80 % of the dataset was used to train the ML models and 20 % of the dataset was used to test the models.

III. RESULTS AND DISCUSSION

Tables II and III summarize the SBP and DBP estimation results of the both datasets for the aforementioned ML approaches. In the tables, the columns are labeled for evaluation according to the AAMI standard and the BHS standard. The best results are highlighted in bold.

TABLE II
RESULTS OF BP ESTIMATION ALGORITHMS ACCORDING TO AAMI AND BHS STANDARD ON DATASET D1_90.

Methods	**BP**	**AAMI [mmHg]**		**BHS [%]**			**BHS**
		MAE	**SD**	**≤ 5, 10, 15 mmHg**			**Grade**
LRM	SBP	7.411	±7.348	48.4	74.6	88.9	C
	DBP	5.034	±5.293	64.1	87.8	95.7	A
k-NN	SBP	6.621	±6.935	54.1	79.8	90.3	B
	DBP	4.666	±5.528	70.5	89.5	95.6	A
DT	SBP	6.792	±7.332	52.1	79.6	90.0	B
	DBP	4.587	±5.431	69.9	90.3	95.5	A
RF	SBP	**5.874**	±5.710	**56.0**	82.6	93.5	B
	DBP	3.983	±4.641	**75.3**	**92.6**	96.5	A
LRMR	SBP	7.372	±7.329	48.6	75.1	89.0	C
	DBP	5.021	±5.296	64.4	87.8	95.7	A
SVR	SBP	7.271	±7.798	51.3	76.2	90.0	B
	DBP	4.911	±5.656	67.0	88.3	95.6	A
AdaBoost	SBP	5.887	**±5.481**	55.7	**83.4**	**93.8**	B
	DBP	4.025	**±4.432**	74.7	92.5	**97.1**	A
Bagging	SBP	5.939	±5.744	55.1	82.4	93.4	B
	DBP	**3.971**	±4.539	75.2	92.4	96.9	A
ANN	SBP	6.688	±6.553	51.2	78.4	90.3	B
	DBP	4.417	±4.891	70.5	90.6	96.4	A
CNN	SBP	7.503	±7.488	48.2	73.4	88.3	C
	DBP	6.053	±5.619	52.8	82.8	94.4	B

TABLE III
RESULTS OF BP ESTIMATION ALGORITHMS ACCORDING TO AAMI AND BHS STANDARD ON DATASET D2_99.

Methods	**BP**	**AAMI [mmHg]**		**BHS [%]**			**BHS**
		MAE	**SD**	**≤ 5, 10, 15 mmHg**			**Grade**
LRM	SBP	4.919	±4.162	61.7	86.9	97.8	A
	DBP	3.069	±3.219	82.0	96.9	99.6	A
k-NN	SBP	5.824	±6.192	60.8	82.0	94.1	B
	DBP	3.112	±3.642	80.6	95.1	99.1	A
DT	SBP	6.056	±4.556	47.8	82.0	95.5	B
	DBP	4.157	±4.186	69.8	94.1	97.8	A
RF	SBP	**4.428**	±3.988	63.5	88.7	98.2	A
	DBP	2.760	±3.088	85.1	**98.2**	99.6	A
LRMR	SBP	4.842	±4.044	60.8	86.9	97.8	A
	DBP	2.947	±3.137	84.2	97.2	99.6	A
SVR	SBP	4.917	±3.997	62.2	87.4	97.8	A
	DBP	2.979	±3.173	82.9	97.2	99.6	A
AdaBoost	SBP	4.494	±3.812	63.9	**90.1**	98.2	A
	DBP	2.772	±3.126	**86.9**	97.8	99.1	A
Bagging	SBP	4.532	**±3.760**	**64.0**	89.6	**99.1**	A
	DBP	**2.738**	**±3.032**	86.0	97.8	**99.7**	A
ANN	SBP	4.891	±4.144	61.3	88.3	97.3	A
	DBP	3.217	±3.339	79.3	96.9	98.7	A
CNN	SBP	6.277	±6.110	55.9	77.5	90.1	B
	DBP	4.232	±4.672	73.9	91.0	96.0	A

Both European and American standards exist to assess whether the BP estimation approach is valid. Here are the two most common assessment standards [7]:

a) BHS standard: In fact, the BHS standard rates pressure measurement systems according to the cumulative percentage of errors below three different thresholds, i.e. 5, 10 and 15 mmHg. Based on this, the BP estimation approach is rated grade A (60 %, 85 % and 95 %), B (50 %, 75 % and 90 %) or C (40 %, 65 % and 85 %) [7].

b) AAMI standard: The BP estimation method meets the standard if the MAE is within 5 mmHg and the standard deviation (SD) is within 8 mmHg [7].

Although more pulse waves were included in D1_90 (N=8084) than in D2_99 (N=1107), only less than 20 % of them are of high quality. This is reflected in the estimation of the BP results, which depend on the quality of the selected PPG waves as shown in Tables II and III. The D1_90 dataset is more general (larger range of BP values), therefore Table II shows more realistic results than Table III.

According to AAMI and BHS, all models for both datasets produce a lower estimation error for DBP.

IV. CONCLUSIONS

There is no significant difference in BP estimation results between the compared ML approaches. The best model on both test sets was the RF method alone or the use of RFs as weaker regressors in AdaBoost or Bagging.

The results of the Adaboost model with RF on the D1_90 test data set were 5.887±5.481 mmHg for SBP and 4.025±4.432 mmHg for DBP (AAMI). The results of the Bagging with RF model on the D2_99 test data set were 4.532±3.760 mmHg for SBP and 2.738±3.032 mmHg for DBP (AAMI). According to the BHS, all 4 estimates for the D2_99 dataset were classified as grade A.

However, the dataset cover only 13 subjects who enter both the training and test sets, and therefore the ML model is overfitted. In future work, we plan to measure more subjects to make the dataset bigger and more variable. Moreover, we plan to use combination of PPG and PCG signal to estimate BP even better.

BP measurement using the smartphone is promising for the future, as it is a cuffless, portable, non-invasive, comfortable and widespread.

REFERENCES

[1] Cardiovascular diseases (CVDs). World Health Organization, 2021, Available from URL: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)/.

[2] S. G. Khalid, J. Zhang, F. Chen and D. Zheng, "Blood Pressure Estimation Using Photoplethysmography Only: Comparison between Different Machine Learning Approaches," Journal of Healthcare Engineering, vol. 2018, Article ID 1548647, 2018.

[3] Telecommunications: Number of smartphone subscriptions worldwide from 2016 to 2021, with forecasts from 2022 to 2027. Statista, 2021, Available from URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/.

[4] G. Slapnicar, N. Mlakar and M. Lustrek, "Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network," Ljubljana, Slovenia, Sensors 19, no. 15: 3420, 2019.

[5] K. Matsumura, P. Rolfe, S. Toda and T. Yamakoshi, "Cuffless blood pressure estimation using only a smartphone," Sci Rep 8, 7298, 2018.

[6] Y. Kurylyak, F. Lamonaca and D. Grimaldi, "A Neural Network-based method for continuous blood pressure estimation from a PPG signal," 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Minneapolis, USA, pp.280-283, 2013.

[7] G. S. Stergiou, B. Alpert, S. Mieke, et al., "A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement," J Hypertens, vol. 36(3), pp.472-478,2018.

Comparison of Pharmacokinetic Models for Quantification of Blood-Brain-Barrier Opening Induced by Focused Ultrasound

Denisa Hývlová
Department of Biomedical Engineering
FEEC, Brno University of Technology
and
Institute of Scientific Instruments
Czech Academy of Sciences
Brno, Czech Republic
xhyvlo00@vut.cz

Jiří Vitouš
Institute of Scientific Instruments
Czech Academy of Sciences
and
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
vitous@isibrno.cz

Radovan Jiřík
Institute of Scientific Instruments
Czech Academy of Sciences
Brno, Czech Republic
jirik@isibrno.cz

***Abstract*—This paper focuses on quantification of blood-brain-barrier (BBB) opening induced by focused ultrasound. Dynamic contrast-enhanced MRI (DCE-MRI) with two different pharmacokinetic models, 2CX and ETK, is used to evaluate BBB opening. The ETK model is commonly used in BBB-opening studies. We propose use of a more advanced model, 2CX, that provides more accurate estimation of the BBB permeability. The accuracy of the models is compared using simulated and real data. The results demonstrate similar performance of the 2CX and ETK models in the brain region with a highly permeable BBB, but the ETK has proven unable to correctly detect low level of permeabilization. Thus, the 2CX model is more suitable for the studies of BBB-permeability dynamics.**

***Index Terms*—DCE-MRI, perfusion analysis, simulation, focused ultrasound, BBB opening**

I. Introduction

The blood-brain barrier (BBB) is a semipermeable membrane that controls the transport of molecules between blood and brain tissue. This membrane protects the brain from harmful substances but also prevents entry of many therapeutic agents. Focused ultrasound (FUS) allows transient, noninvasive and local opening of the BBB, and thus, targeted delivery of drugs. [1] This technique has shown promise for the treatment of a variety of brain disorders, including Alzheimer's disease, Parkinson's disease, and brain tumors [2] [3].

FUS is applied after intravenous administration of a microbubble contrast agent (CA). Acoustic waves cause oscillation of microbubbles in the vasculature, which results in disruption of tight endothelial junctions of the BBB, thus increasing its permeability. This effect persists for several hours, up to 24 hours [1].

This work was supported by Czech Science Foundation (GA22-10953S); computational infrastructure and MR instrumentation of the ISI-MR facility of the Czech-BioImaging infrastructure was funded by grants LM2018129 and LM2023050 of the MEYS CR.

Although there are reported clinical studies using FUS-induced BBB opening for drug delivery [2] [3], preclinical studies gauging safety, efficiency and precision prevail. The level of BBB opening might be evaluated by dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI). This technique is based on capturing T1-weighted images before, during and after the intravenous administration of a gadolinium-based CA. BBB opening enables the passage of particles with diameters up to 65 nm; thus magnetic resonance (MR) CAs with sizes typically between 1 and 60 nm [1] are suitable for permeabilization evaluation.

Passage of the CA results in contrast enhancement in images corresponding to concentration of the CA in the tissue at the given time $c(t)$, which can be modeled as

$$c(t) = I(t) * c_a(t), \tag{1}$$

convolution of the impulse residue function $I(t)$ and the CA concentration in plasma of the feeding artery $c_a(t)$, also known as the arterial input function (AIF). The time curves $c(t)$ and $c_a(t)$ can be measured, the unknown of this equation is only the signal $I(t)$, which can be obtained by fitting of (1) to the measured concentration curves [4]. The parameters of the model of $I(t)$ are the perfusion parameters, such as plasma flow F_p, plasma volume v_p, extravascular extracellular volume v_e or permeability-surface product PS, depending on the chosen pharmacokinetic (PK) model. In FUS experiments, PS is the most important parameter to be estimated as it quantifies the level of BBB opening.

FUS-induced BBB-opening studies using DCE-MRI to assess the permeabilization of the BBB usually employ the Tofts model [5] [6] or more commonly the extended Tofts model [7]–[9], often denoted as the extended Tofts-Kety (ETK) model. The ETK model is defined by the following impulse residue function:

$$I(t) = v_p\delta(t) + K^{\text{trans}} e^{-tK^{\text{trans}}/v_e}. \tag{2}$$

DOI 10.13164/eeict.2023.133

Therefore, the primary perfusion parameters derived from this model are v_p, v_e and K^{trans}. Here, K^{trans} reflects the influx of the CA from plasma into the interstitium and is used as a measure of permeabilization in BBB-opening studies. The problem is that K^{trans} is a perfusion parameter that includes the effects of both F_p and PS. Only in the permeability-limitted case (high F_p and low but nonzero PS), K^{trans} is assumed to be approximately equal to PS [4].

In contrast to the ETK, selection of a so called advanced PK model, that describes the CA pharmacokinetics more realistically (at the cost of higher complexity and parametrization by more perfusion parameters), allows separation of F_p and PS from K^{trans} and should represent the BBB opening more accurately. One of the eligible advanced PK models is the two-compartment exchange (2CX) model, with impulse residue function

$$I(t) = F_p(Ae^{\alpha t} + (1 - A)e^{\beta t}), \tag{3}$$

where A, α and β are model parameters that are directly used to compute v_p, v_e and PS [4]. To the authors' knowledge, this model has not been applied for quantification of BBB opening in FUS applications so far. One reason for this might be high demands of advanced PK models on temporal resolution of the DCE-MRI acquisition. Here, we are overcoming this problem by using compressed-sensing approach, including radial golden-angle acquisition and spatially-regularized image reconstruction [10].

In this study, we assess the accuracy of the permeabilization estimates using the commonly used ETK model and a the 2CX model. First, the models are compared using simulated DCE-MRI data, which allows comparison to the ground truth. The observed results are then illustrated using a real dataset.

II. Methods

A simulator designed at the Institute of Scientific Instruments of the Czech Academy of Sciences was used to generate the synthetic DCE-MRI data. The input for the simulation is a tissue-structure rat phantom composed of segmented tissue regions in an axial slice of the head. Each tissue was assigned perfusion parameters based on literature or on real perfusion-analysis results, parametrization of the PK models was chosen as $\{F_p, E, v_e, T_c\}$, where E is the extraction fraction and T_c the mean capillary transit time. The simulator generates the concentration curves using a predefined PK model and AIF for each region of interest and converts it to signal intensity (SI, intensity in the MR image) based on fast low-angle shot (FLASH) acquisition. A synthetic image sequence with high temporal and spatial resolution is then created from the SI curves and the tissue-structure phantom, and weighted by coil sensitivities supplied as an input. MR echo signals are then generated by transformation of the synthetic image sequence to the k-space. The simulator allows both Cartesian and radial sampling of the k-space.

Perfusion analysis of both synthetic and real data was performed using the PerfLab software [11] designed for quantitative DCE-MRI analysis. This software performs voxel-by-voxel perfusion analysis using various PK models, providing estimates of perfusion parameters.

A. Concentration curves analysis

To evaluate the performance of the chosen PK models under ideal conditions and confirm their suitability for this application, perfusion analysis based only on simulated concentration curves was performed. The signals were generated for various values of PS that may be induced by FUS. Other perfusion parameters were considered unaffected by sonication and kept at constant values expected in the healthy brain, i.e. F_p = 0.8 ml/min/ml, v_e = 0.1 ml/ml and T_c = 0.01 min. The value of E was calculated using the equation for compartment models [4]

$$E = \frac{PS}{F_p + PS}. \tag{4}$$

PS was sampled in the interval $< 0, 0.2 >$ ml/min/ml with a step of 0.01 ml/min/ml. This range covers the expected values that might be induced by FUS, which could be up to 0.06 ml/min/ml [5] [6], and also values induced after exceeding the safe FUS parameters.

The concentration curves were modeled using the 2CX model (3), and then subsampled to achieve four different sampling periods of the signals, $T_s = \{0.5, 1, 2, 3\}$ s. Perfusion analysis of the simulated curves using the PerfLab software with the ETK and 2CX models was performed to obtain K^{trans} and PS estimates, respectively.

B. Synthetic-dataset analysis

The next step involved simulation and analysis of a complete synthetic dataset. First, the input tissue-structure data of the simulation software had to be modified to accurately represent the FUS-induced BBB opening. Therefore, the FUS region was added to the tissue-structure phantom (Fig. 1 left) with 4 levels of permeabilization defined by $PS = \{0.2, 0.4, 0.5, 0.6\}$ ml/min/ml. These values were assigned based on our previous perfusion analysis of a real DCE-MRI dataset obtained shortly after the FUS application. Similar to the concentration curve analysis, the extraction fraction E was calculated according to (4) and the remaining perfusion parameters were considered to be unaffected by FUS.

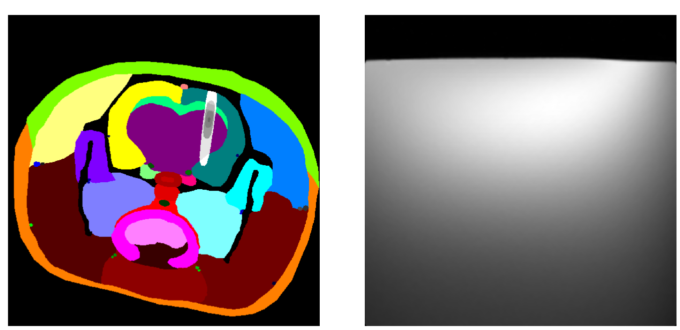


Fig. 1. Tissue-structure rat phantom (left), gray segments correspond to the added FUS region. Measured sensitivity of the FUS coil (right).

Another modification to the current simulation software resided in the sensitivity estimation of the coil used in the FUS experiments. This was performed by measuring MR images using a homogeneous phantom. The most representative slice near the center of the coil was selected from the obtained data. The image was adjusted according to the tissue-structure phantom geometry and parameters, resulting in the coil sensitivity estimate shown in Fig. 1 right.

The parameters of the simulated 2D radial golden angle data were based on a possible real acquisition: repetition time TR = 8 ms, echo time TE = 1.445 ms, flip angle FA = 15°, 38400 radial projections and matrix size of the reconstructed images 128×128. Precontrast multi-FA scans were simulated with the parameters FA=$\{1, 3, 5, 10, 15, 20, 25, 30\}°$ and 2880 projections. DCE-MRI image sequences were reconstructed from the simulated k-space dataset using the BART reconstruction toolbox [12]. The reconstruction was performed with L1 wavelet spatial regularization and temporal total-variation regularization, with weights set appropriately according to our previous experience. The temporal resolution was set to 0.768 seconds and the resulting image series was processed using the PerfLab software with the ETK and 2CX models.

C. Real-dataset analysis

The results obtained from the synthetic-data analysis were validated by processing a real dataset of BBB opening induced in a rat (approved by the National Animal Research Authority). The application of FUS was immediately followed by MR-CA administration and measurement of DCE-MRI data on a 9.4T Bruker BioSpec MRI GmbH preclinical scanner. Radial 3D golden angle stack-of-stars acquisition [10] was used with the following parameters: TR = 8 ms, TE = 1.475 ms, FA = 15°, 40000 projections and matrix size of the reconstructed images 128×128×8. The sampling period of the reconstructed data was 0.96 seconds and the dynamic sequence was processed using the PerfLab software with the ETK and 2CX models.

III. Results

A. Concentration curves analysis

A comparison of the permeabilization estimates of both models with the ground truth (GT) PS for four different sampling periods is shown in Fig. 2. The results show different behaviors of the models for higher and lower values of PS. For higher PS, K^{trans} estimates deviate from the GT values of PS by underestimating them. This inaccuracy seems to be temporal sampling indepedent for the ETK model. The PS estimates of the 2CX model deviate from the GT with increasing PS to a larger extent than for the ETK model, and in a different direction (overestimating them). This overestimation is more severe with increasing sampling period. The effect of the sampling period is known from literature as the problem of discretized convolution integral [13].

Although the ETK model outperforms the 2CX model in this range of permeability values, safe FUS-induced permeabilization would be lower than approximately 0.06 ml/min/ml [6], hereby not so significantly affected by this phenomenon.

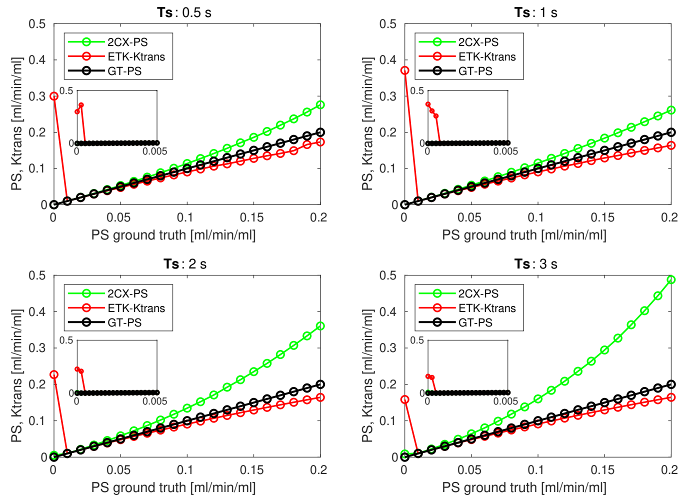


Fig. 2. Estimated permeability by the 2CX model (green) and the ETK model (red) compared to the ground truth (black) for various induced permeability levels and various sampling periods of the simulated signals.

In the range of low PS values induced by FUS, both models show higher accuracy in their permeability estimates. The 2CX model fits accurately over the entire range of BBB opening values with less apparent influence of the sampling period. On the contrary, the ETK model failed to estimate the absence of permeabilization, and this inacurracy was increasing with higher temporal resolution. Based on this observation, a finer sampling of low PS values was realized to see the model's behavior for close-to-zero PS, with values in the range of $< 0, 0.005 >$ ml/min/ml with a step of 0.00025 ml/min/ml, and the experiment was repeated for these values. The results are shown in the insets of the graphs in Fig. 2. The ETK model showed this inaccuracy also for nonzero PS GT values, whereas the 2CX model fitted exactly the GT.

B. Synthetic-dataset analysis

The results of the synthetic-dataset perfusion analysis using both models compared with GT PS are shown in a close-up view of the phantom in Fig. 3. The 2CX model provides an overall accurate estimation of PS for all tissues compared with the GT. A minor overestimation is apparent in the FUS area. K^{trans} given by the ETK model significantly underestimates vessel permeabilization level for most tissues due to the influence of blood flow as the permeability-limitted conditions

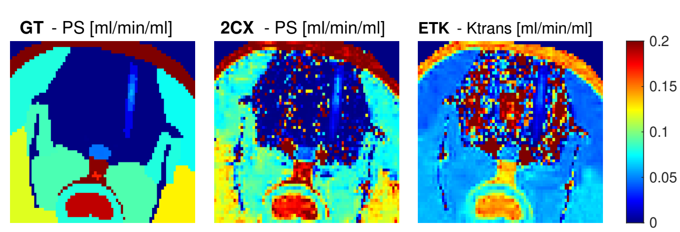


Fig. 3. Perfusion maps of the ground truth PS, the 2CX model estimate of PS and the ETK model estimate of K^{trans} for the synthetic dataset.

TABLE I
MEAN ABSOLUTE ERROR OF THE PK MODELS' PERMEABILITY ESTIMATES IN THE FUS AND BRAIN REGIONS.

MAE	*2CX*	*ETK*
FUS	0.0100 ml/min/ml	0.0084 ml/min/ml
Brain	0.0615 ml/min/ml	0.1333 ml/min/ml

of the ETK model do not hold outside of the FUS region. A slight underestimation was also observed in the FUS region.

Considering only the brain, which is the region of interest in the BBB-opening study, the obtained results confirmed the phenomena observed in the previous analysis (Fig. 2). For the FUS region and unaffected brain, the mean absolute error (MAE) was determined. The results are summarized in Table I. In the area of BBB opening, the 2CX model showed slightly worse performance with MAE = 0.01 ml/min/ml compared to 0.0084 ml/min/ml in the case of ETK. The difference between the models was more significant for the brain tissue unaffected by sonication, where the real PS is zero because of the intact BBB. In this area, the ETK model showed more than two times higher MAE than the 2CX model, with 0.1333 ml/min/ml compared to 0.0615 ml/min/ml.

C. Real-dataset analysis

The results of the perfusion analysis on a real dataset are shown in a close-up view of the brain in Fig. 4. The obtained perfusion maps correspond with perfusion analysis executed on the synthetic dataset, that is, significant overestimation of the ETK model in the brain with an intact BBB. The differences between the models in the level of BBB opening in FUS-exposed region have proven negligible, mainly owing to acquisition and reconstruction noise. The ETK model shows a slightly larger BBB opening area in the K^{trans} map. This might be an effect of the model's overestimation observed for small levels of permeabilization, as shown in Fig. 2.

IV. CONCLUSION

This study focused on comparing the commonly used ETK model quantifying the level of FUS-induced BBB opening with the K^{trans} rate constant and an advanced PK model, 2CX, separating plasma flow F_p and permeability-surface product PS, thus allowing a more accurate description of permeabilization.

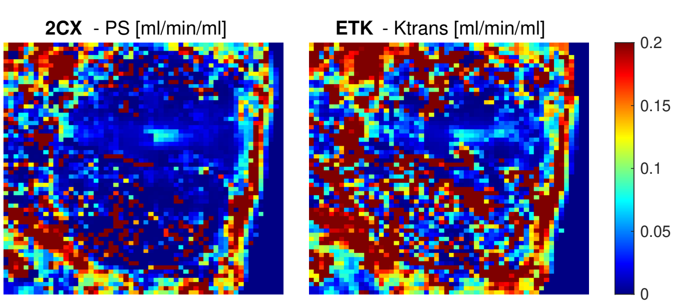


Fig. 4. Perfusion maps of the 2CX model estimate of PS and the ETK model estimate K^{trans} for the real dataset.

In the range of PS values induced by FUS, both models showed accurate estimations with ETK performing slightly more accurately. Real-dataset analysis confirmed the results; both models provided similar estimations of the BBB opening level, and the differences between the 2CX and ETK models were negligible.

The major difference between these models appeared for an intact BBB with zero PS and for small, close-to-zero values of PS. Perfusion analysis with the ETK model resulted in a significant overestimation of permeabilization and provided an incorrect estimation of PS for the majority of brain tissue. In contrast, the 2CX model estimation was mainly accurate, with only a few false estimates. These mistakes are most likely caused by incorrect local extrema found during optimization in PK modeling.

The false estimates produced by the ETK model in regions with low vessel permeabilization might lead to misinterpretation of the results obtained during BBB closure. Thus, the 2CX model has proven to be more suitable for evaluation of the BBB-closure dynamics.

REFERENCES

[1] A. Conti, H. A. Kamimura, A. Novell, A. Duggento, and N. Toschi, "Magnetic resonance methods for focused ultrasound-induced blood-brain barrier opening," *Frontiers in Physics*, vol. 8, 2020.

[2] N. Lipsman, Y. Meng, A. J. Bethune *et al.*, "Blood–brain barrier opening in Alzheimer's disease using MR-guided focused ultrasound," *Nature Communications*, vol. 9, 2018.

[3] Y. Meng, B. J. MacIntosh, Z. Shirzadi *et al.*, "Resting state functional connectivity changes after mr-guided focused ultrasound mediated blood-brain barrier opening in patients with Alzheimer's disease," *NeuroImage*, vol. 200, pp. 275–280, 2019.

[4] S. P. Sourbron and D. L. Buckley, "Classic models for dynamic contrast-enhanced MRI," *NMR in Biomedicine*, vol. 26, pp. 1004–1027, 2013.

[5] T. Sun, G. Samiotaki, S. Wang *et al.*, "Acoustic cavitation-based monitoring of the reversibility and permeability of ultrasound-induced blood-brain barrier opening," *Physics in Medicine & Biology*, vol. 60, pp. 9079–9094, 2015.

[6] F. Vlachos, Y. S. Tung, and E. Konofagou, "Permeability dependence study of the focused ultrasound-induced blood-brain barrier opening at distinct pressures and microbubble diameters using DCE-MRI," *Magnetic Resonance in Medicine*, vol. 66, pp. 821–830, 2011.

[7] C. H. Fan, W. H. Lin, C. Y. Ting *et al.*, "Contrast-enhanced ultrasound imaging for the detection of focused ultrasound-induced blood-brain barrier opening," *Theranostics*, vol. 4, pp. 1014–1025, 2014.

[8] P. C. Chu, W. Y. Chai, C. H. Tsai *et al.*, "Focused ultrasound-induced blood-brain barrier opening: Association with mechanical index and cavitation index analyzed by dynamic contrast-enhanced magnetic-resonance imaging," *Scientific Reports*, vol. 6, 2016.

[9] K. Yoon, W. Lee, E. Chen *et al.*, "Localized blood-brain barrier opening in ovine model using image-guided transcranial focused ultrasound," *Ultrasound in Medicine & Biology*, vol. 45, pp. 2391–2404, 2019.

[10] L. Feng, R. Grimm, K. T. Block *et al.*, "Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI," *Magnetic Resonance in Medicine*, vol. 72, no. 3, pp. 707–717, 2014.

[11] R. Jiřík, "Perfusion lab." [Online]. Available: http://perflab.cerit-sc.cz/

[12] M. Uecker, "BART: Computational magnetic resonance imaging." [Online]. Available: https://mrirecon.github.io/bart/

[13] T. S. Koh, D. L. H. Cheong, and Z. Hou, "Issues of discontinuity in the impulse residue function for deconvolution analysis of dynamic contrast-enhanced MRI data," *Magnetic Resonance in Medicine*, vol. 66, no. 3, pp. 886–892, 2011.

Estimation of blood glucose level based on PPG signals measured by smart devices

Enikö Vargová
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech republic
xvargo02@vutbr.cz

Andrea Němcová
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech republic
nemcovaa@vut.cz

***Abstract* — This paper deals with the possibilities of non-invasive determination of blood glucose from photoplethysmographic signals. Monitoring blood sugar is the most important part of managing diabetes. Diabetes is one of the world's major chronic diseases. Untreated diabetes is often a cause of death.**

Two datasets have been created by recording the photoplethysmographic signals of 16 people using two smart devices (a smart wristband and a smartphone), along with their blood glucose levels measured in an invasive way. The photoplethysmographic signals were preprocessed, and suitable features were extracted from them. The aim of the work is to propose methods for glycemic classification and prediction.

Various machine-learning models were created. The best model for classifying blood glucose into two groups (low blood glucose and high blood glucose) is random forest, which achieves an F1 score of 84% and 80% on two different test sets obtained from two smart devices. The best blood glucose level prediction model is also based on random forest and achieves an MAE of 1.02 mmol/l and 1.17 mmol/l on both testing datasets.

***Keywords* — PPG, glycemia, diabetes, smartphone, smart devices, classification, prediction**

I. Introduction

Diabetes is one of the world's major chronic diseases. The cause of this disease is a relative or absolute lack of insulin. The International Diabetes Federation (IDF) reports that approximately 537 million people in the world, aged 20–79 years, suffer from diabetes. Untreated diabetes is often a cause of death. [1]

Blood glucose monitoring is critical for diabetics to keep their blood glucose levels (BGL) within the target range. In the current clinical field, only an invasive blood glucose detection method is used. However, it has many disadvantages. Patients suffer from physical and mental pain, and there is a high risk of infection in this invasive method. [2]

For these reasons, this paper focuses on the possibility of determining BGL non-invasively from the photoplethysmographic (PPG) signal obtained using smart devices such as a smart wristband (Empatica) or a smartphone. The use of smart devices in healthcare is cheap, user-friendly, on-the-spot, and suitable for long-term self-monitoring at home. It allows remote sharing of measurement records and results with medical doctors.

II. Materials and Methods

A. Data Recording

For the purpose of this experiment, two databases of PPG signals were created. One database contains data measured by the Empatica wristband. The other database contains PPG signals recorded by a smartphone (Samsung Galaxy Note 20).

Sixteen subjects willingly took part in the experiment (11 female and 5 male, ages 22 to 79), of whom 5 were diagnosed with diabetes or prediabetes. Informed written consent was obtained from all probands. The measurement was approved by the Faculty of Electrical Engineering and Communication ethics committee.

The measurement scheme is in Fig. 1. Each measurement lasted 15 minutes. The PPG signal was recorded by Emaptica wristband throughout the measurement period. During the measurement, PPG signal of one minute duration was recorded three times using the smartphone, as shown in Fig. 1.

Smartphone PPG is extracted from video. Video is recorded from the subject's fingertip by the smartphone's rear camera. The fingertip is illuminated by camera LED flash to enhance the quality of the signal. At the same time, the reference blood glucose value was measured. It was measured three times, as shown in Fig. 1. For this purpose, certified glucometer Fora Diamond Mini was used. Probands diagnosed with diabetes were measured twice.

B. Preprocessing

To obtain more data for classification and prediction, the 15-minute PPG signals recorded by the Empatica wristband were divided into three 5-min segments, one for each of the three reference blood glucose levels. If the reference value is greater than the chosen threshold (*7.2 mmol/l*), the 5-min segment is reduced to 4.5 min. Otherwise, it was shortened to 2.5 min. Subsequently, the segments were divided into 10-second segments (Fig. 1). The threshold value was determined according to the manual of the glucometer, which displays a specific blood glucose value and also classifies the given value according to the threshold of 7.2 mmol/l into normal blood glucose and elevated blood glucose.

The video measured using the smartphone camera was converted to a PPG signal. The red video channel was selected, and then each video frame was averaged to get a single value. The PPG signals obtained from the smartphone were also divided into 10-second segments.

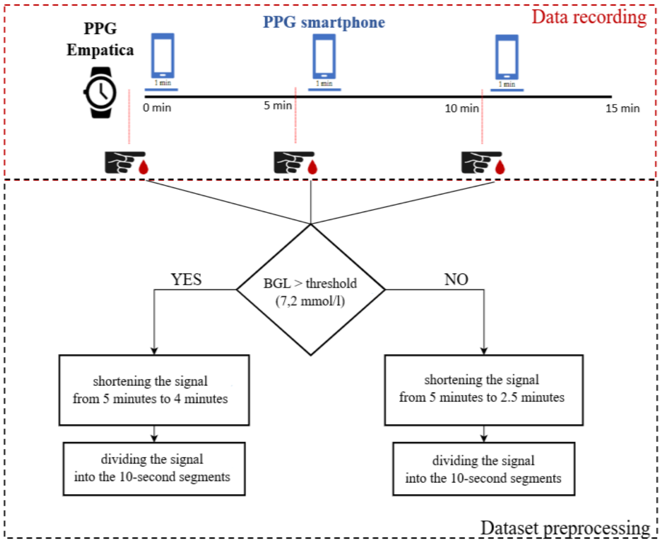


Fig. 1. Scheme of data measurement and dataset preprocessing

Low-quality PPG signals measured by the smartphone were removed using an algorithm proposed in [5]. However, some poor-quality signals were not removed by the algorithm, thus manual quality annotation was additionally performed.

All 10-second signals were inverted, filtered using a second-order Butterworth bandpass filter (0.6–3.6 Hz), and then max-min normalization was performed (Fig. 2).

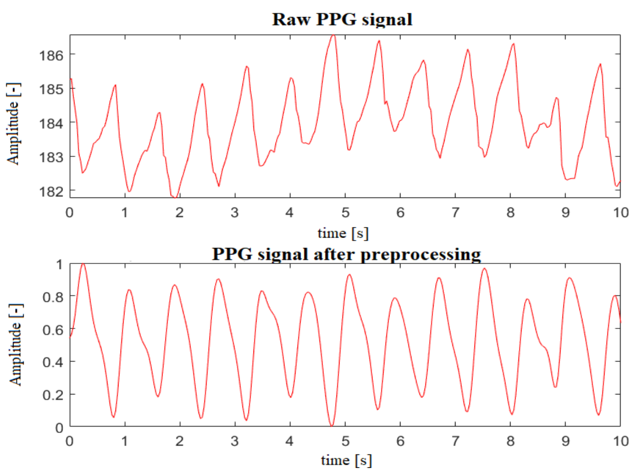


Fig. 2. PPG signal obtained from a smartphone before and after preprocessing

C. Feature extraction

Twenty-seven different features were extracted from the PPG signals. Some features are based on heart rate variability (HRV) analysis, others on the morphology of the signal itself.

Features were standardized using the Z-score, outliers were replaced, and statistical analysis of features was performed:

- Kolmogorov-Smirnov test (test of normality)
- Mann-Whitney U test (nonparametric test)
- Unpaired t-test

The Kolmogorov-Smirnov test is used to test the null hypothesis that a dataset comes from a normal distribution. If the feature has a normal distribution, an unpaired t-test was used for further analysis. Otherwise, the Mann–Whitney U test was used. The Mann–Whitney U test and the unpaired t-test are statistical tests that determine whether there is a difference between two unrelated groups (low and high blood glucose). As a result of these tests, 12 features were excluded from further analysis. Then the Sequential Forward Selection (SFS) method was used to select the best combination of features using a Random Forest (RF) classifier. The method chose 9 features (TABLE I.). These features were used for both classification and prediction tasks.

TABLE I. DESCRIPTION OF SELECTED FEATURES

Feature	Description
kurtosis	describes the distribution of a dataset around the mean and tells us whether the data is distributed symmetrically or not
skewness	is a measurement of the distortion of symmetrical distribution or asymmetry in a data set
peak to RMS	returns the ratio of the largest absolute value in signal to the root-mean-square (RMS) value of signal
RMSSD	root mean square of successive differences between normal heartbeats
average PP interval	average time interval between two adjacent PPG peaks from a 10-s PPG signal
SD1/SD2	ratio of short-term variability of PP intervals SD1 and long-term variability of PP intervals SD2
IQR	the interquartile range of the signal that shows the difference between the 75th quantile and the 25th quantile
shortest PP interval	shortest distance of PP intervals
standard deviation of the 1st derivative of the PPG signal	standard deviation of the 1st derivative of the PPG signal

D. Train-Test split

The aim of this paper is to propose methods for glycemic classification and prediction. Both training and testing took place on the dataset from the Empatica smart wristband. Data from different subjects are used for training and testing. The proposed algorithms were subsequently verified on a completely independent dataset containing data from the smartphone. The division of data into training and testing sets and the representation of data in individual groups are shown in TABLE II.

TABLE II. TRAINING AND TESTING SETS

	Number of 10-s records	**Low glycemic level**	**High glycemic level**
Training set (Empatica)	784	400	384
Testing set (Empatica)	341	149	192
Testing set (smartphone)	129	60	69

The BGL in the datasets were in the range of: 4.3–13.2 mmol/l (Training set), 5.2–9.7 mmol/l (Testing Set, Empatica) and 4.3–13.2 mmol/l (Testing set, smartphone).

E. Classification of glycemia

The first main goal of this paper is to classify PPG signals into two groups – low glycemic level and high glycemic level (threshold *7.2 mmol/l*). For this purpose, selected features are used. Three models are proposed for the classification task – RF classifier (number of estimators = 60, minimum sample leaf = 17), Support Vector Machines (SVM) classifier (kernel = polynomial), and K-Nearest Neighbors (KNN) classifier (k = 11).

F. Prediction of blood glucose level

The second main goal of this paper is blood glucose level prediction. Three models were proposed to predict blood glucose from the PPG signal – RF regressor (number of estimators = 50, minimum sample leaf = 3), SVM (kernel = Radial Base Function (RBF)), and KNN (k = 11). The same data sets as for the classification task were used for both training and testing.

III. RESULTS

As can be seen from TABLE III, the classification models achieve very similar results. Accuracy and F1 scores are lower on the smartphone dataset than on the data from the Empatica wristband. This fact was assumed since the entire dataset comes from another device and was not used in the training phase.

TABLE III. CLASSIFICATION RESULTS

Classification	RF		SVM		KNN	
Set	***Acc***	***F1***	***Acc***	***F1***	***Acc***	***F1***
Training set (Empatica)	87%	87%	82%	84%	84%	84%
Testing set (Empatica)	82%	84%	82%	83%	82%	84%
Testing set (smartphone)	75%	80%	76%	80%	78%	82%

Fig. 3 shows the area under the receiver operating characteristic (ROC) curve for individual models on the testing set from the Empatica device, according to which RF is the best model for classification.

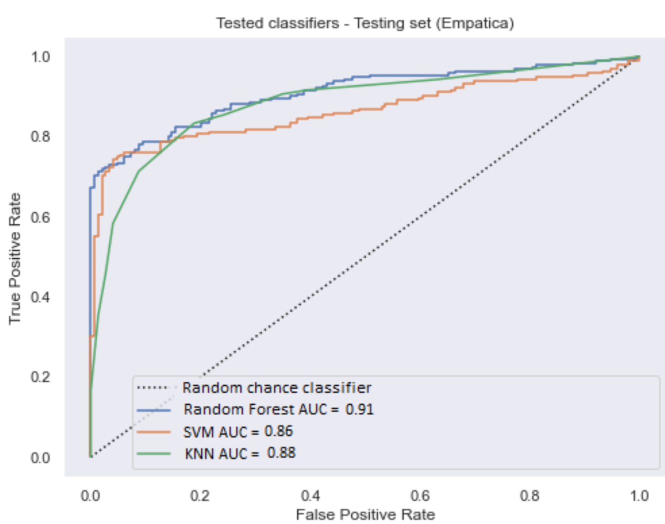


Fig. 3. AUCs for the testing data obtained from the Empatica wristband

However, on the smartphone testing set, RF and KNN models performed equally well according to the area under the ROC curve (Fig. 4). The SVM model performed slightly worse on both testing sets.

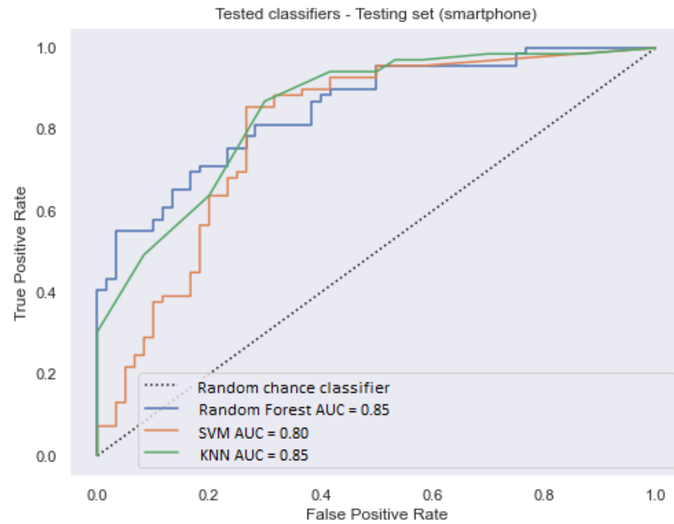


Fig. 4. AUCs for the testing data obtained from the smartphone

Furthermore, models for predicting BGL were evaluated. The mean absolute error (MAE) for individual models is in TABLE IV. RF achieves the best results.

TABLE IV. PREDICTION RESULTS

Prediction	RF	SVR	KNN
Set	***MAE [mmol/l]***	***MAE [mmol/l]***	***MAE [mmol/l]***
Training set (Empatica)	0.48	0.85	0.83
Testing set (Empatica)	1.02	1.64	1.33
Testing set (smartphone)	1.17	1.66	1.37

IV. DISCUSSION

The proposed RF models for both classification and regression tasks achieve results comparable to those of other authors. Zhang et al. [4] used PPG signals acquired from smartphones for the classification of the glucose level (normal, borderline, and warning) with an accuracy of over 80%. Islam et al. [3] created a model that can predict the actual glucose level with an error of less than 20 mg/L (1.11 mmol/l). The effectiveness of the proposed models to classify and predict BGL can be limited by the number of measured subjects and the range of the measured BLG (4.3–13.2 mmol/l). However, the results of this experiment confirm the usefulness of the smart device based noninvasive estimation of BGL.

V. CONCLUSION

The paper deals with the classification and prediction of blood glucose level from PPG signals obtained from two smart devices – Empatica wristband and smartphone (Samsung Galaxy Note 20). The measured data were preprocessed, and appropriate features were extracted from them. Subsequently, various models were created for both classification and prediction tasks.

The best model for classifying blood glucose into two groups (low blood glucose and high blood glucose) is RF (number of estimators = 60, minimum sample leaf = 17). It achieves an F1 of score 84% and 80% on both testing sets (the testing set from Empatica and the testing set from the smartphone).

The best blood glucose level prediction model is based on RF (number of estimators = 50, minimum sample leaf = 3). It achieves an MAE of 1.02 mmol/l and 1.17 mmol/l on both testing datasets.

The proposed RF models for both classification and regression tasks achieve results comparable to those of other authors. The results of this experiment confirm the usefulness of the smart device based noninvasive estimation of BGL.

REFERENCES

[1] H. Deng, L. Zhang, Y. Xie, and S. Mo, "Research on estimation of blood glucose based on PPG and deep neural networks," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 693, no. 1, p. 012046, 2021.

[2] "IDF diabetes atlas," *Diabetesatlas.org*. [Online]. Available: https://www.diabetesatlas.org.

[3] T. T. Islam, M. S. Ahmed, M. Hassanuzzaman, S. A. Bin Amir, and T. Rahman, "Blood glucose level regression for smartphone PPG signals using machine learning," *Appl. Sci. (Basel)*, vol. 11, no. 2, p. 618, 2021

[4] G. Zhang *et al.*, "A noninvasive blood glucose monitoring system based on smartphone PPG signal processing and machine learning," *IEEE Trans. Industr. Inform.*, vol. 16, no. 11, pp. 7209–7218, 2020.

[5] E. Vargová, "Stanovení kvality a odhad tepové frekvence ze signálu PPG," Bakalářská práce. Vysoké učení technické, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Andrea Němcová, pp. 51, Brno, 2021.

Design of a Three Phase Fault Tolerant Nested Neutral Point Piloted Converter with Active Capacitor Voltage Balancing and Space-Vector Modulation

Bc. Dominik Klement
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
xkleme13@vutbr.cz

Dr. Jun Wang
Power & Energy Systems Laboratory
University of Nebraska-Lincoln
Lincoln, United States
junwang@unl.edu

Ing. Jiří Janoušek
DTEE
Brno University of Technology
Brno, Czech Republic
xjanou09@vutbr.cz

This paper presents a design for a nested neutral point piloted (NNPP) converter. The proposed design incorporates redundant control of each power stage and additional sensing and control circuitry required to implement a fault tolerant operation. Additionally, an active capacitor voltage balancing technique is proposed and implemented to enhance the converter's volumetric density and improve the output harmonics by mitigating capacitor voltage imbalance. Simulation results validate the effectiveness of the proposed design. Hardware prototype is designed and is currently being assembled.

I. INTRODUCTION

Multilevel converters offer advantages over conventional two-level converters in terms of output voltage quality, such as lower harmonic distortion and over voltage values on the output terminals. However, each topology has its own advantages and disadvantages. While diode-clamped converters are simple and reliable, they require many clamping diodes. Flying capacitor converters are less complex but suffer from voltage balancing issues and when a cascading design is used also a number of capacitors and more complex control strategy. Comparison of frequently used converter topologies is shown in Table I.

TABLE I. COMPARISON OF CONVERTER TOPOLOGIES [1]

	Two-level Converter	Three-Level Converter	Five-level SMC
Output Voltage THD	71.9 %	43.3 %	**26.3 %**
Output Current THD	15.6 %	10.0 %	**3.8 %**
Max Over-Voltage Value	2300 V	1700 V	**1300 V**

Stacked Multicell Converter topology is theoretically infinitely scalable without major control algorithm modifications, the control algorithm developed as a part of this research paper allows for N-level converter to be implemented, where 13-level converter was simulated.

II. CONVERTER DESIGN

A. Selecting a Topology

A Stacked Multicell Topology as shown in Fig. 1. was selected due to being infinitely scalable without major modifications to the power stage nor the algorithm. This topology consists of a cell that makes it possible to share the voltage constraints on several switches, making it ideal for use in medium and high voltage inverters with wide bandgap semiconductors.

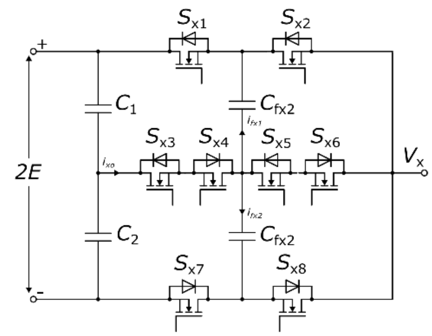


Fig. 1. Five-Level Stacked Multicell Converter topology

B. Design of the control algorithm

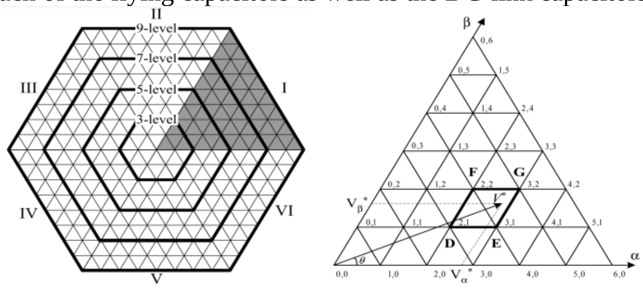
Compared to a conventional three phase two-level converter where a Half-bridge topology is used to control the output voltage, there are, based on the output level switching state analysis up to 729 different possible output combinations [2], some of which are redundant and have the same effect on the output voltage but are crucial due to their effect on the voltage balance on the capacitors. Which in turn have great effect on the output voltage harmonics as well the possible use of the redundant states in fault tolerant algorithms.

TABLE II. SWITCH COMBINATIONS OF A FIVE-LEVEL SMC CONVERTER

S_{x1}	S_{x2}	S_{x3}	S_{x4}	S_{x5}	S_{x6}	S_{x7}	S_{x8}	V_{ox}	I_{fx1}	I_{fx2}	I_{x0}
1	1	1	1	1	0	0	0	E	0	0	0
1	0	1	1	1	1	0	0	$E/2$	$-I_{ox}$	0	0
0	1	1	1	1	0	0	0	$E/2$	I_{ox}	0	I_{ox}
1	0	1	0	0	1	0	1	0	$-I_{ox}$	$-I_{ox}$	0
0	0	1	1	1	1	0	1	0	0	0	I_{ox}
0	1	0	1	1	0	1	0	0	I_{ox}	I_{ox}	0
0	1	0	1	1	0	1	0	$-E/2$	0	I_{ox}	0
0	0	0	1	1	1	1	0	$-E/2$	0	$-I_{ox}$	I_{ox}
0	0	1	1	0	0	0	1	$-E$	0	0	0

Table II. shows the effect of each of the switch combinations on the output voltage and capacitor current.

Selection of the voltage space vector that is used in a switching sequence is based on a function which takes into account the output voltage, based on which the magnitude of the output level (in this case five levels, but can theoretically expanded as shown in Fig. 2.) is selected, when the selected level is redundant, the algorithm also takes into account the voltage of each of the flying capacitors as well as the DC-link capacitors.



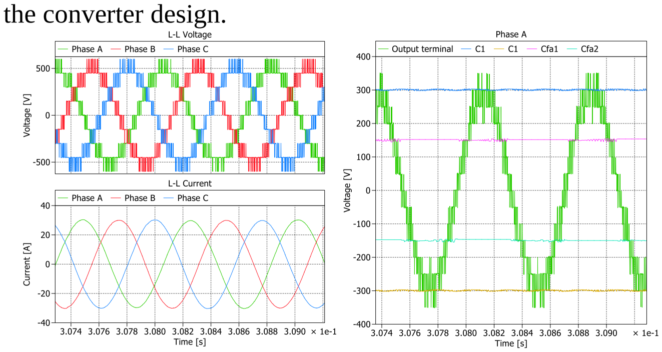
(a) (b)

Fig. 2 Space vectors (a) for a 9-level converter (b) in a single 60° sector [3]

The use of Space Vector Modulation (SVM) on the converter power stage allows the output voltage to be as much as 15.15 % greater when compared to a tradition Sinusoidal Pulse Width Modulation (SPWM). [4] Further details about this algorithm are beyond the scope of this paper.

C. Simulation of the converter

For the proposed control algorithm to be developed a simulation model of the converter with approximate parameters was designed. PLECS by Plexim simulation platform was chosen due to the powerful tools available. Such as the capability to perform Hardware in Loop tests with the real converter power stage and the ability to comprehensively analyze the output. Fig. 3. Shows an output of the simulation of the converter design.



(a) (b)

Fig. 3. Output of the prototype simulation (a) Line to Line (b) Phase A

The same software along with a component library was used to estimate the switching and conduction losses of the converter during operation. This information was then used to calculate the heat dissipation solution for both the power switches and capacitors.

III. Prototype design

A. Design parameter selection

The selection of the components used for the real-world tests of the converter were based on requirements which were derived from the possible use cases of such converter i.e., medium voltage high speed motor drives and high voltage direct transmission line converters. The voltage and power used in a prototype was limited to 600 volts and 15 kilowatts respectively. This was done due to safety reasons and costs associated with high voltage components as well as the availability of a suitable power supply and a load for testing.

TABLE III. Overview of the converter design parameters

Converter Parameter	Value
DC-link voltage	600 V
Power rating	15 000 W
DC-link capacitance	140 μF
Floating capacitance	20 μF
Load resistance	5 Ω
Load inductance	1.8 mH
Fundamental frequency	1300 Hz
Switching frequency	50 000 Hz

Based on requirements shown in Table 3. and budget considerations the main components were selected. The power switch used is a Silicon Carbide (SiC) N-Channel Power MOSFET Microsemi MSC035SMA070B4. Each phase of the converter contains 16 power switches, where every switch is isolated and controlled individually. Each switch also contains a voltage and current measurement circuitry, which is to be used as an input for the fault tolerant operation algorithms. Both the DC-link as well as the flying capacitors have a voltage rating of 500 volts DC, such a high value has been chosen to ensure safe operation even if an error occurs during the testing of active capacitor voltage balancing algorithms.

B. Power board design

First an analysis of the commutation loop of the converter was performed, this ensures the minimization of the effects of electromagnetic interference as well as the possibility of voltage overshoot and the inherent damage of the components. The schematic as well as the layout of the power stage, which is shown in Fig. 4. of the converter including a separate gate driving circuitry was done in Altium Designer 22. Block design approach was used, which allows for a simplified future expansion of the stacked levels as well as the possibility to use any of the blocks in a different converter topology.

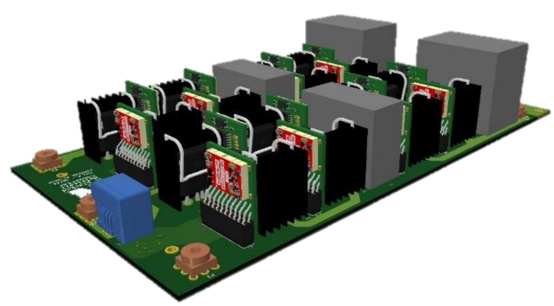


Fig. 4. Rendering of a single-phase power stage board

C. Control board design

The algorithm is to be run on a digital signal processor (DSP) from Texas Instruments, specifically DSP on the control card TMDSCNCD28379D shown in Fig. 5. for which a daughter board is designed. The limited number of input and output capabilities of the DSP only allow for operation of the converter using a reduced PWM control scheme, which is sufficient for the purposes of tests of the converter hardware design, but not sufficient for full fault tolerant capabilities, in the future a FPGA will be used for this purpose. Such a controller was chosen due to the ease of integration as well as native support of the PLECS Processor in Loop (PiL) module, which allows for rapid testing process.

Fig. 5. Texas Instruments TMDSCNCD28379D

The control board is to be connected to the power stage using IDC cable, the design also allows for connection with DSPACE controller.

D. Sic MOSFET Driver board module

Due to the limited area on the power board caused by the need to minimize the commutation loop, there is not enough space for a MOSFET driver with the necessary additional circuitry, such as a dual isolated power supply, gate driver and gate driving resistors.

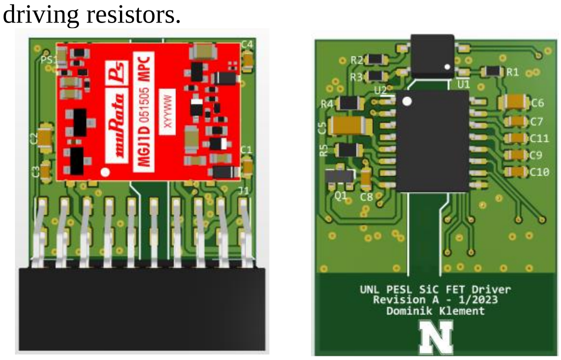


Fig. 6 Rendering of the SiC Gate driver module

A separate board shown in Fig. 6. is designed. It is then connected to the main power board via pin headers. This driver board can be used in future revisions of the converter, saving resources, and accelerating prototyping.

IV. FAULT TOLERANT PROPOSAL

Due to the increased complexity in terms of number of switches used in a stacked multicell topology, the converter is inherently less reliable when using a control algorithm without a fault tolerant operation implemented. However, with a fault tolerant algorithm used, the converter is very robust and can resist situations, which would make a tradition two level topology stop operating. This property is particularly useful in devices of high importance, where a malfunction could result in endangerment of lives, such as loss of propulsion of an e-plane powertrain or extreme financial losses, such as a downtime of a high-power transmission line or a fault in a satellite antenna pointing system motor controller.

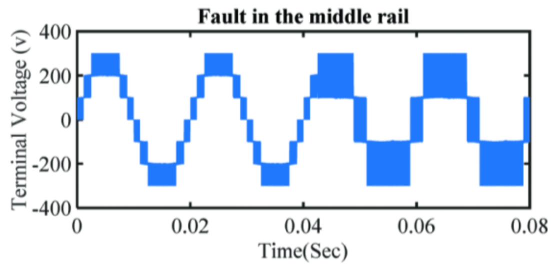


Fig. 7 Fault tolerant algorithm output waveform [5]

The algorithm uses window comparators to detect a fault in the circuit. This allows for a fast response and modification of the control strategy. Fig. 7. shows an example of a response to a fault on a seven level SMC converter, fault occurred at the time of 44 milliseconds.

V. CONCLUSION

This paper presents a design for a Nested Neutral Point Piloted Converter with Active Capacitor Voltage Balancing. The proposed design features redundant control of each power stage, an active capacitor voltage balancing technique, and additional sensing and control circuitry to implement a fault-tolerant operation. The simulation results validate the functionality of the proposed algorithm, a hardware prototype is currently being assembled. The Stacked Multicell Topology is selected for its scalability, making it ideal for use in medium and high voltage inverters with wide bandgap semiconductors. The control algorithm developed allows for N-level converters to be implemented, with a 13-level converter being simulated. The proposed design improves output harmonics, enhances volumetric density, and mitigates capacitor voltage imbalance, making it a promising solution for medium and high voltage applications.

REFERENCES

[1] J. Yu, R. Burgos, Q. Wang, and I. Agirman, “Design of a SiC-based Five-Level Stacked Multicell Converter for High-Speed Motor Drives,” in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, Sep. 2019, pp. 4063–4068. doi: 10.1109/ECCE.2019.8912728.

[2] Q. Cheng, C. Wang, X. You, Z. Chen, and Z. Dong, “Voltage-Balancing Control in Stacked Multicell Converter Based on Single Space-Vector Modulation,” *IEEE Trans. Power Electron.*, vol. 37, no. 1, pp. 344–353, Jan. 2022, doi: 10.1109/TPEL.2021.3098843.

[3] S. Wei, B. Wu, F. Li, and C. Liu, “A general space vector PWM control algorithm for multilevel inverters,” in *Eighteenth Annual IEEE Applied Power Electronics Conference and Exposition, 2003. APEC ’03.*, Feb. 2003, vol. 1, pp. 562–568 vol.1. doi: 10.1109/APEC.2003.1179268.

[4] F. Sevilmiş and H. Karaca, “Simulation and analysis of SVPWM based VSI for wind energy systems,” in *Proceedings of the 2014 6th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Oct. 2014, pp. 73–78. doi: 10.1109/ECAI.2014.7090151.

[5] P. Hekmati, I. P. Brown, and Z. J. Shen, “Fast Detection of Open Circuit Device Faults and Fault Tolerant Operation of Stacked Multilevel Converters,” in *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*, Mar. 2019, pp. 1758–1765. doi: 10.1109/APEC.2019.8721923.

University of Nebraska-Lincoln Power & Energy Systems Laboratory

Indoor Monitoring System For Well-being

1st Jakub Kočka
Department of Control and Instrumentation
Brno University of Technology
Brno, Czechia
xkocka03@vutbr.cz

2nd Jakub Arm
Department of Control and Instrumentation
Brno University of Technology
Brno, Czechia
Jakub.Arm@vut.cz

***Abstract*—The work address the indoor air quality monitoring enabling to monitor the well-being and comfort of occupants. A monitoring system is able to measure room temperature, humidity, barometric pressure, CO2, luminosity, and dust pollution using chip sensors. The measurement devices are connected wirelessly via ZigBee 3.0 technology enabling mesh network topology and the measured data is collected into a database. The data will be then processed and presented using a cloud application enabling advanced control and alarm management, when a value exceeds the health limits. The measurement system comprises of measurement devices and a gateway. The measurement device is designed as a sandwich architecture comprising sensor and communication module. The created sensor module is interfaced with Adafruit Feather interface in order to use any available Adafruit SoC kit as the communication module. The conducted results also suggest that ESP32-H2, which is now in pre-production phase, operates as ZigBee 3.0 SoC in an embedded device.**

***Index Terms*—Indoor air quality, ZigBee 3.0, ESP32-H2, NRF52840, Smart Home, IoT**

I. Introduction

Indoor air quality (IAQ) is an issue on which many companies and health-care researcher are focusing to improve user's health and well-being in a space. The purpose of a monitoring system is to provide data based on biological, physical, and chemical sensors to improve maintenance and management strategies enabling also other benefits like energy reduction. [12]

IAQ monitoring system is traditionally designed as a star topology system comprising of measuring devices and edge device (gateway) transmitting data to a cloud service. The communication could be achieved by a local area network technology; in order to reduce the power consumption of a device, ZigBee technology (moreover, offers mesh network topology) could be selected. [13]

To integrate the ZigBee technology in a MCU-based device, SoC chips ensure small footprint while sufficient computational power. Among the leading market chips, nRF52840 and CC2530 could be highlighted. Other manufactures try to enter the market, such as Espressif with ESP32-H2 chip, which is now in the pre-production phase. Another way implementing ZigBee into a device rests in use of XBee module or its derivatives [14].

The purpose of the work is to create a measurement system for various room variables using COTS (Commercial off-the-shelf) devices to collect IAQ data. Another purpose rests in the validation of the use of ESP32-H2 (as the begin to market SoC) and ZigBee 3.0 technology. The sensor module was designed and created. The communication module firmware was also designed and created in order to validate the proper selection. After it is integrated into a system comprising cloud database and web data application, the measurements will be conducted and the system could be validated. Due to the validation of the system architecture, database and visualization were implemented only locally and tested separately.

Our work contributes several innovations, which other indoor monitoring systems could benefit from. Our measurement device architecture allows to combine different sensor and communication module which follow Adafruit Feather interface; thus, a commercially available communication module is able operate with our sensor board. Moreover, the sensor board could be replaced to measure a different set of variables. We could also highlight the use of dust pollution and luminosity sensor. Another outcome of the work rests in the one of the early applications of ESP32-H2 as the ZigBee communication SoC.

II. System architecture

The system can be divided into four parts. There are two hardware parts (sensor unit and communication module based on SoC), which create together the whole measurement device as it is shown at the figure 1, and two software parts (database and visualization).

The measurement devices (acting as ZigBee end devices) are periodically awaken from sleep mode to perform a measurement; then, the data is transmitted via a ZigBee network into the gateway (Windows OS based machine acting as ZigBee coordinator) and stored in the database running on the machine. The same machine hosts visualization which scrapes data from the database.

The used database is the InfluxDB enabling time series storage architecture which is suitable for periodic measurement collection. As a visualization software, a lot of commercial and open-source software can be used, which InfluxDB 2.0 or Grafana application could be suitable candidates from. Grafana has been selected to favor micro-services architecture against a centralized solution. In the integration phase of the project, the cloud option of database and visualization will be adopted.

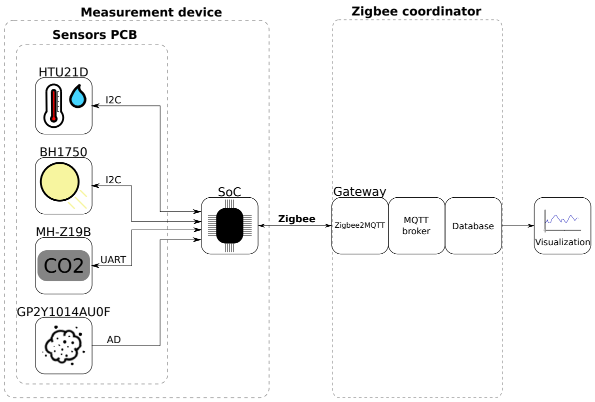


Fig. 1. Block diagram of the monitoring system

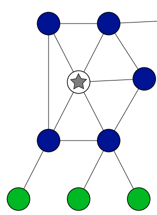


Fig. 2. ZigBee mesh topology

A. ZigBee

ZigBee is wireless communication protocol invented in 1998 by ZigBee Alliance (In nowadays CSA - Connectivity Standards Alliance). This communication standard is designed to have low consumption requirements in battery powered devices. The standard has been built on top of the wireless standard IEEE 802.15.4 for low-rate wireless personal area networks.

The ZigBee usage enables a mesh topology which is shown at figure 2. In this topology there are devices connected between each other except of the end devices that makes the topology self-healing which means that even when the link to the device is down, the new different route to the device can be established if there exists one.

In ZigBee network, devices can play three roles:

- coordinator (star - Fig. 2, 3) - manages the network and runs an application
- routers (blue - Fig. 2, 3)) - makes routes to target devices and runs an application
- end devices (green - Fig. 2, 3)) - runs an application

1) Link Quality Indicator: Link Quality Indicator (LQI) is characterization of strength and quality of received packet.

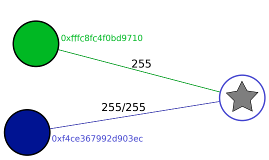


Fig. 3. ZigBee test network with maximal LQI (real distances = 1m)

This measurement should by done for every received packet ([10], page 63). This indicator can acquire 0 to 255.

The figure 3 shows small test network which has been created with nRF52840 SoC (blue) acting as a router device and ESP32-H2 (green) acting as a end device. Above the links to coordinator there is displayed LQI, because of the short distances (1 m each, without obstacles) the LQI reached maximal value 255.

2) Security: Security of a ZigBee network is based on link and network keys. The keys have size of 128 bits. Link keys are for unicast communication and broadcast communication uses network key shared with every device in the network. ([11], page 400)

III. Design of Measurement Device

This chapter describes design of measurement device. Embedded sensors enable temperature, humidity, barometric pressure, CO2, luminosity, and dust pollution measurement. Communication between sensors and SoC is via I2C, UART, or AD converter.

The Communication Module and the Sensor Module are designed to have compatible bus with each other and they are Adafruit Feather compatible. This design makes measurement device modular in built-in sensors and in data transfer used technology. Figure 4 shows the design and integration of the sensor module.

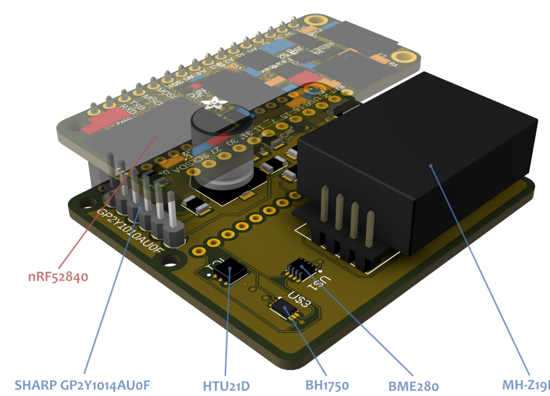


Fig. 4. Sensor module visualization

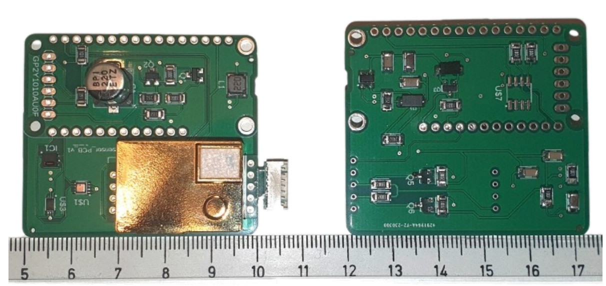


Fig. 5. The created sensor module (units in cm)

A. Communication Module

The communication module contains SoC and components with power management and there should be for example USB to UART bridge for direct firmware upload from computer. Also there should be battery power management with battery level monitor circuit.

The figure 3 shows connection of different SoCs used in ZigBee network. The blue node in the picture is nRF5240 acting as a ZigBee router and the green node is a ESP32-H2 acting as a ZigBee end device.

B. Sensor Module

The sensor module is designed and created to measure lot of variables while small dimensions are achieved. The module is powered from battery, which has power management on the communication Module. On Sensor Module board there is step-up for 5 V so there are 3.3 V and 5 V busses which means both voltage level sensor types can be used.

For communication between sensors and SoC there are two busses - I2C and UART. Also AD converter pin and PWM is embedded if needed in another combinations of sensor communication module.

Although level-shifters and other converters need to be used, the created sensor module gives a reference sensing module that is compatible with COTS communication modules.

IV. Indoor environment recommendations

This chapter summarizes recommendations for indoor environment from different standards for variables that are being monitored by the system. This recommendations can be used in analysis, for alerts, or as a target value of actuators where the actual value is given by this monitoring system.

A. Temperature and humidity

The temperature and humidity are the most common comfort parameters which can be easily determined by a human perception.

Thermal comfort can be determined and interpreted of thermal comfort using calculation of PMV (predicted mean vote). In European standard ISO 7730 [1] there could be found for example the operative temperature for single office should be 24.5 $\pm$ 1 $°C$ at summer and 22 $\pm$ 1 $°C$ at winter.

B. Light intensity

Another important parameter for well-being is the light intensity. For specific activities, there is a need for certain level of the light intensity; the values are described in Table I.

C. Carbon Dioxide

Carbon dioxide (CO2) can be in high concentration for human even mortal invisible gas. So it is good to adjust it's level in the room at reasonable concentration. The concentrations are in Table II.

TABLE I
TABLE LIGHT INTENSITY AND SPECIFIC ACTIVITIES ([2], PAGE 355)

Type of the space, task and activity	***Light intensity [lx]***
archive	200
reception	300
document storing, copying and so on	300
writing, reading, data analysis	500
CAD working station	500
conference room	500
technical drawing	750

TABLE II
POTENTIAL DANGER OF CO2 CONCENTRATION [3]

Description	***CO2 [ppm]***
average outdoor concentration	400
typical indoor concentration with good air flow	400 – 100
poor air quality	1000 – 2000
headache, sleepiness, lost of concentration	2000 – 5000
health problems	5000
immediate life danger in lack of oxygen	>40000

D. Dust

Breathing a dust particles can lead to respiration issues. Limits are described in Table III.

V. Sensor device parameters

Following chapter describes used sensors and relations between the real world value and measured value.

A. Temperature and humidity

The temperature and relative humidity sensor is measured by I2C SMT chip HTU21D with parameters: [5]

- 3.3 V
- humidity 12 bit resolution
- temperature 14 bit resolution
- humidity: 0–100 %, temperature: -40 – 125 $°C$ range

B. Light intensity

Light intensity is measured by the SMT chip based on I2C communication BH1750 with parameters: [6]

- 3.3 V
- humidity 16 bit resolution
- 0–65535 lx range

TABLE III
DUST PARTICLES LIMITS FOR SAFETY OF HUMAN LIFE [4]

Particle size [μm]	***Average time***	***Limit [$\mu g \cdot m^3$]***
10	24 h	50
10	1 year	40
2.5	24 h	25
2.5	3 years	20

C. Carbon dioxide

CO2 sensor embodies a MH-Z19B chip. It uses non-dispersive infrared (NDIR) principle for measurement. The output options are UART which provides direct CO2 concentration or PWM which needs a calculation to get CO2 concentration. Chip parameters: [7]

- 5 V
- 5000 ppm range
- accuracy $\pm$(50 ppm + 5 $\%$ reading value) ppm

D. Dust

Dust particles sensor uses optical principle for measurement. It is the SHARP GP2Y1014AU0F sensor. The dust density is given by the linear dependency of output voltage. Chip parameters: [8]

- 5 V
- 0.5 mg/m^3 range
- accuracy 0.5 $V/(0.1mg/m^3)$

E. Barometric pressure

Sensor for barometric pressure is BME280 which combines pressure, humidity and temperature sensor. The additional temperature and humidity sensor allows more accurate measurement for this values: [9]

- 3.3 V
- relative humidity
  - 0 – 100 $\%$ range
  - accuracy $\pm$3 $\%$
- pressure
  - 300 – 1100 hPa range
  - accuracy $\pm$1.7 hPa

VI. Future vision

When the system is implemented in a real environment, the data will be gathered to form a dataset. From the dataset, a model could be created to analyze the air properties in a room while daytime and season are changed; a preliminary research was conducted by [15].

The data and model could be helpful to estimate properties of an air quality for room appliances, such as an air conditioning system (HVAC), which could be then controlled based on the data in an advanced way to reduce the energy consumption; a preliminary research was conducted by [16].

Conclusion

The article presents the design of an indoor air quality monitoring system for well-being, which should alert in case of any of the monitored value is out of the comfortable limit. The crucial part of the system is a measurement devices, which provides data to the server via a gateway using ZigBee 3.0 technology enabling mesh network. The measurement device is designed to provide a possibility to combine commercially available communication module with the created sensor module via Adafruit Feather interface. Besides the standard monitored variables, the sensor module is equipped by chips to measure luminosity and dust pollution. The results also suggest that ESP32-H2 communication module, which is now in the pre-production phase, is suitable for ZigBee 3.0 embedded devices.

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] Ergonomics of the thermal environment — Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria. 2005.

[2] HABEL, Jiří. Světlo a osvětlování. Praha: FCC Public, 2013. ISBN 978-80-86534-21-3.

[3] Carbon Dioxide [online]. Wisconsin: Department of Health Services, 2022 [cit. 2023-03-09]. Available at: https://www.dhs.wisconsin.gov/chemical/carbondioxide.htm

[4] PETRLÍK, RNDr. Jindřich Petrlík, Ing. Petr Válek VÁLEK a Ing. Milan Havel HAVEL. Polétavý prach - PM [online]. Praha: Arnika, 2010, 2022 [cit. 2023-03-09]. Available at: https://arnika.org/toxicke-latky/databaze-latek/poletavy-prach-pm

[5] HTU21D(F) RH/T SENSOR IC: Digital Relative Humidity sensor with Temperature output [online]. Mouser Electronics: Measurement Specialties, Inc., a TE Connectivity company, 2017 [cit. 2023-03-09]. Available at: https://cz.mouser.com/datasheet/2/418/5/NG_DS_HPC199_6_A1-1128627.pdf

[6] Digital 16bit Serial Output Type Ambient Light Sensor IC: BH1750FVI [online]. Mouser Electronics: ROHM Co., 2011 [cit. 2023-03-09]. Available at: https://components101.com/sites/default/files/component_datasheet/BH1750.pdf

[7] Infrared CO2 Sensor Module: (Model: MH-Z19B) [online]. https://www.winsen-sensor.com/: Zhengzhou Winsen Electronics Technology Co., 2019 [cit. 2023-03-09]. Available at: https://www.winsen-sensor.com/d/files/MH-Z19B.pdf

[8] GP2Y1014AU0F: Compact Optical Dust Sensor [online]. Philippines: Sharp [cit. 2023-04-01]. Available at: https://global.sharp/products/device/lineup/data/pdf/datasheet/gp2y1010au_e.pdf

[9] BME280: Combinated humidity and pressure sensor [online]. Germany: Bosch, 2022 [cit. 2023-04-01]. Available at: https://datasheet.lcsc.com/lcsc/1811141121_Bosch-Sensortec-BME280_C92489.pdf

[10] IEEE Std 802.15.4-2003: Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs) [online]. New York: The Institute of Electrical and Electronics Engineers, 2003 [cit. 2023-03-09]. Available at: https://people.iith.ac.in/tbr/teaching/docs/802.15.4-2003.pdf

[11] ZigBee Specification [online]. ZigBee Alliance: ZigBee Alliance, 2015 [cit. 2023-03-09]. Available at: https://zigbeealliance.org/wp-content/uploads/2019/11/docs-05-3474-21-0csg-zigbee-specification.pdf

[12] Liu, X.M., Wang, F.F., Zeng, Z.B. Design and Implementation of Indoor Environmental Quality Monitoring System Based on ZigBee. Proceedings of the International Conference on Computer Information Systems and Industrial Applications (CISIA). ATLANTIS PRESS, 2015.

[13] Liu, Z., Wang, G., Zhao, L. a Yang, G. Multi-Points Indoor Air Quality Monitoring Based on Internet of Things. IEEE Access, vol. 9, pp. 70479-70492, 2021. DOI 10.1109/ACCESS.2021.3073681.

[14] Salman, N., Andrew, H., Kemp, A., Khan, A.H., Noakes, C.J. Real Time Wireless Sensor Network (WSN) based Indoor Air Quality Monitoring System. IFAC-PapersOnLine, Volume 52, Issue 24, 2019, Pages 324-327. DOI 10.1016/j.ifacol.2019.12.430.

[15] Marzouk, M., Atef, M. Assessment of Indoor Air Quality in Academic Buildings Using IoT and Deep Learning. Sustainability 2022, 14, 7015. DOI 10.3390/su14127015.

[16] Li, Y., Ma, N., Guo, L. Energy-Constrained Indoor Air Quality Optimization for HVAC System in Smart Building. IEEE Systems Journal, vol. 17, no. 1, pp. 361-370, March 2023. DOI 10.1109/JSYST.2022.3159566.

3D Scanner of Electromagnetic Field

Filip Dusek
Department of Radio Electronics
FEEC, Brno University of Technology
Technicka 12, 616 00 Brno, Czech Republic
xdusek30@vutbr.cz

Jaroslav Lacik
Department of Radio Electronics
FEEC, Brno University of Technology
Technicka 12, 616 00 Brno, Czech Republic
lacik@vut.cz

***Abstract*—This contribution deals with the design of an electromagnetic field 3D scanner and its basic electromagnetic probes. The design of the scanner was done in SOLIDWORSK 2022. The movements of the 3D scanner are realized along the XYZ axes by step motors. To scan the electromagnetic field, E and H probes were designed in ANSYS HFSS. The scanner controlled by a personal computer is able with a spectrum analyzer to sense the electromagnetic field up to 10 GHz.**

Keywords—3D scanner of electromagnetic field, E-probe, H-probe, near-field measurement

I. Introduction

A 3D scanner of electromagnetic (EM) field is used to measure radiating near field of a device under test (DUT). On the market, two types of devices are emerging. One device involves measurement using a robotic arm (manufacturer: Sensitivity in Frequency APREL type EM-ISight-4 [1]), while the other device looks similar to the 3D printer (manufacturer: EMAG TECHNOLOGIES INC. type NEOSCAN® [2]).

The basic requirements for the presented 3D scanner are the following:

- the ability to measure devices up to the size of A4,
- the probe proximity to the measured device at 1 mm,
- the weight of the tested object should not exceed 1.5 kg,
- the scanner's design should be reconfigurable,
- the ability to measure from several kHz up to 10 GHz,
- minimal impact of the scanner's construction on electromagnetic measurement,
- a cheap yet robust implementation.

Due to these requirements, the 3D scanner concept of the 3D printer was chosen for implementation.

Based on the above requirements the 3D EM fields is designed for measuring DUT by a spectrum analyzer or vector network analyzer. The scanner is controlled by a personal computer (PC). After the measurement, the software for the scanner is able to depict the measured field by 3D graphs and the measured data can be further processed using relevant software. The 3D graphs can provide better image of the near-field radiation. For the measurement by the scanner, E-field and H-field coaxial probes were designed in ANSYS HFSS. They are based on the open ended or loop ended coaxial line.

II. Configuration of 3D EM Field Scanner

The design of the scanner was done in SOLIDWORKS 2022 with the aim to be cost-effective and reconfigurable, with minimal influence on the measured results. The proposed construction of the scanner is depicted in Fig. 1. The scanner composes from several basic parts:

- Aluminum profiles (drawn by grey in Fig. 1): They assure stability and reconfigurability of the scanner.
- Step motors (drawn by red in Fig. 1): They assure spinning of the trapezoidal thread, which allow movement of the placement plate, elevator, and platform.
- Guide rods and bearings (drawn by grey in Fig. 1): They assure stability and minimize friction during the movement.
- Other parts (drawn by green in Fig. 1) they assure the correctness of the functionality, these parts are printed on the 3D printer.

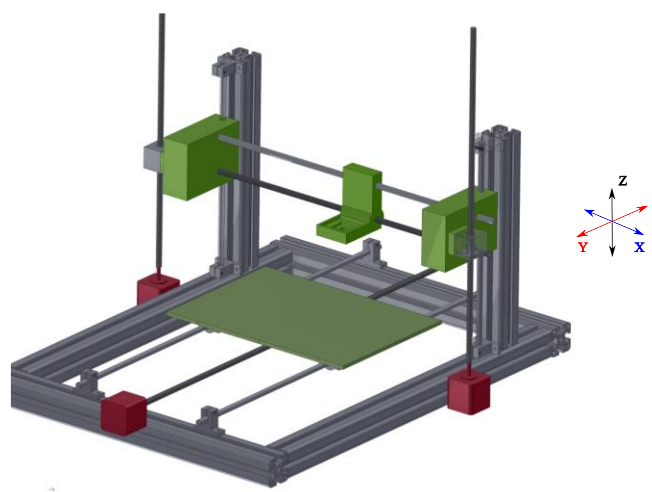


Fig. 1. Configuration of the proposed 3D EM field scanner.

The basic design of the scanner includes four step motors: the one for moving the placement plate in the Y axis, two for moving the probe in the Z axis to approach or move away from

This work was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-23-8191.

the placement plate, and one for moving the measuring platform in the X direction. There is an option to add a fifth motor to rotate the probe.

The minimum load of the step motors was calculated. However, in order to assure good operation of the step motors, the chosen step motors can be loaded at least three times more that the calculated values. The result of this evaluation is summarized in Table I [4].

TABLE I. CALCULATED AND CHOSEN MOTORS

Name	*Calculated (Nm)*	*Dimension*	*Chosen (Nm)*	*Type*
Placement plate	0.0108	15x	0.17	NEMA 17
Elevator	0.4407	3x	1.24	NEMA 23
Platform	0.0124	13x	0.09	NEMA 17

The principle of the scanning is depicted in Fig. 2. A personal computer controls the scanner and reads the measured data from the spectrum analyzer measured by a probe. In the computer, the measured data are stored and further processed. Three step motors of the scanner allow moving of the probe in the X, Y and Z direction. The probe in XY plane moves in a meander way as depicted in Fig. 3 to scan desired area. The movement of the Z axis allows the probe to move closer or further away from the measured object. The step of the probe movement in the X, Y and Z direction is defined before the scanning by a user.

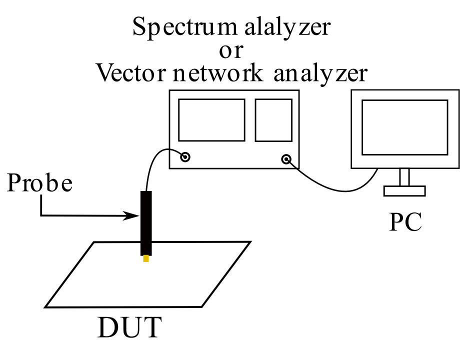


Fig. 2. The principle of measuring electromagnetic fields and storing data.

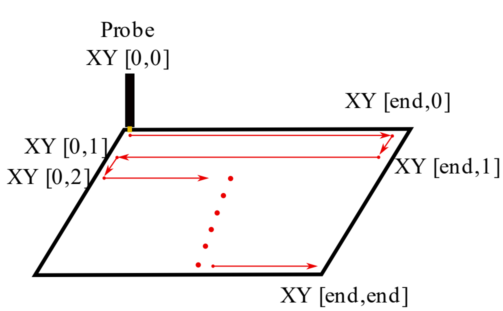


Fig. 3. The principle of moving probe.

The 3D scanner is designed to measure objects of A4 size. The overall dimensions of the scanner are set to be 640 mm in length, 477 mm in width, and 340 mm in height. The total weight of the scanner does not exceed 15 kg.

III. DESIGN OF PROBES

We only get measured results from the near field testing with probes. The test can show where the radiation is stronger and weaker and what type of field is dominant. This gives us the ability to modify the device or isolate the problem and thus solve the problem.

There are two types of probes: electrical field probes and magnetic field probes. The simplest probe to measure the electric field is the open end of the coaxial cable, where we measure the electric field caused by the voltage. [5] They are mainly used for measuring PCB traces leading to high impedance logic. The sensitivity of the open end of coaxial cable as a probe is insufficient, it can detect an electric field of around 5 mV/m.

To measure the magnetic fields, the simplest type is a loop from a coaxial cable. [5], [6] Compared to the electric probe, it is more sensitive to field detection, and is able to detect the current at about 5 µA. We get the maximum output through the probe when the loop is aligned with the leading current and is in close proximity to it. Thanks to the directional properties, the user can trace the electrical path with excessive radiation. Rotating probes can cause large differences in the measured amplitude of the given magnetic field.

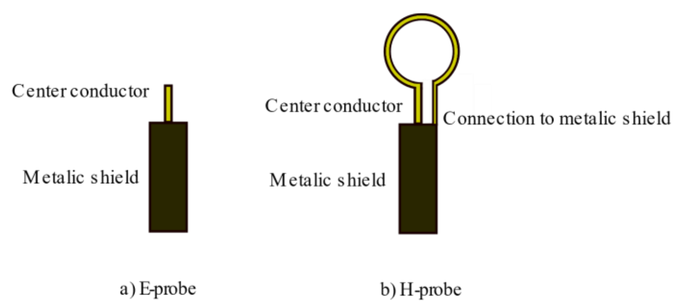


Fig. 4. (a) Electric field probe and (b) magnetic field probe.

Both types of probes were modeled in the ANSYS HFSS. The probes were designed from Semi-rigid coaxial cable, which has the parameters listed in Table II [3]. The length of the probes is 10 cm.

TABLE II. THE PARAMETRS OF THE SEMI-RIGID COAXIAL CABLE

	Inner Conductor	*Isolation*	*Outer Conductor*
Material	SPWC	PTFE	Copper/Tin
Radius	0.46 mm	1.49 mm	1.79 mm

To simulate the transmission coefficient in the ANSYS HFSS, it is necessary to create identical probes next to each other. The distance between the centers of the probes is 1 cm. These basic probes were simulated in the frequency range from 1 kHz to 10 GHz. The simulation results are depicted in the figures Fig. 5 and Fig. 6.

Fig. 5. Transmission coefficient betweem two electric field probes.

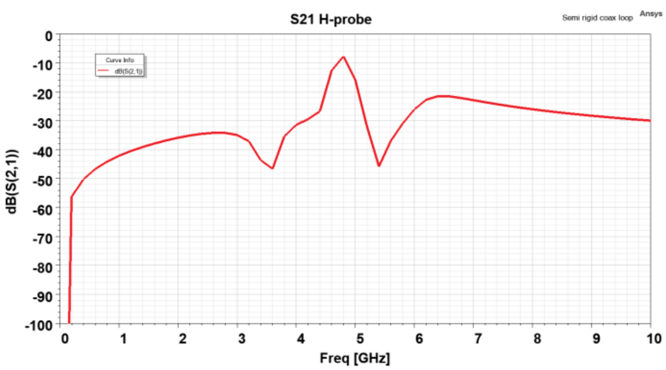


Fig. 6. Transmission coefficient betweem two of magnetic field probes.

IV. Utility Program

The 3D EM scanner is controlled by the PC. A graphical user interface (GUI) was created in the MATLAB environment. In order to control the scanner, it is possible to use the NUCLEO development board from the STM32 company, because MATLAB has the capability to interface with STM32 through an extension module [7]. Motors connected to the NUCLEO are controlled thought a GUI. A spectrum analyzer or vector network analyzer sends data to the PC, and this data can be displayed in 3D graphs in the GUI once the measurement is complete.

The GUI also includes options such as interrupting the measurement, returning to the default position and zooming in/out with the probe.

V. Conclusion

The 3D scanner of electromagnetic field was designed in SOLIDWORKS 2022. The operation principle of the scanner was explained. Motors for the XYZ axis movements were calculated and dimensioned. In addition, the principles of the electromagnetic field measurement and data storage from the spectrum analyzer or the vector network analyzer were explained. Furthermore, the principle of probe movement was described.

Two types of basic measuring probes, the electric field probe and magnetic field probe, were designed and simulated in ANSYS HFSS. The transmission coefficient was simulated in the frequency range from 1 kHz to 10 GHz. The magnetic field probe is more suitable for measuring fields at lower frequencies. In contrast, the electric field probe is more suitable for measuring higher frequencies.

The control program for the 3D scanner will be created using the MATLAB. To control the scanner, it is necessary to connect MATLAB to a PC with the NUCLEO development board from STM32.

The scanner is currently being implemented, and it is expected to be operational during the student conference EEICT.

References

[1] APREL, EM-ISight. *Sensitivity in Frequency APREL*, Accessible on: https://www.aprel.com/em-isight

[2] EMAG TECHNOLOGIES INC, NEOSCAN® *EMAG Technologies INC.*, Accessible on: https://emagtech.com/neoscan-system/#

[3] TAOGLAS®. *CAB.058*, Current Version 17.1.2020, Accessible on: https://www.farnell.com/datasheets/3195865.pdf

[4] MBMTECHNIK. Krokové motory: Hodnoty motorů *Mbmtechnik*, Accessible on: https://www.mbmtechnik.cz/elektromotory/krokove-motory/krokove-motory-36

[5] KRAZ, Vladimir. CREDENCE TECHNOLOGIES, INC.,"*Near-Field Methods of Locatinf EMI Sources"*, Engineering Magazine, pp 1-7 May/June 1995, Accessible on: https://www.ramayes.com/download/OnFILTER/OnFILTER-Near-Field-Methods-of-Locating-EMI-Sources.pdf

[6] OLIVEIRA, Fernando. DIY Near Field Probes & Preamplifiers. *Interference Technology*, November 29, 2018, Accessible on: https://interferencetechnology.com/diy-near-field-probes-preamplifiers/

[7] MathWorks®, "STMicroelectronics Hardware Support from Simulink", Accessible on: https://www.mathworks.com/products/hardware/stm32.html#nucleo-boards

Comprehensive analysis of putrescine metabolism in *A. thaliana* using GWAS, genetic risk score, metabolic modelling and data mining

Jana Schwarzerová
Department of Biomedical Engineering
Faculty of Eletrical Engineering and Communication,
Brno University of Technology
Brno, Czech Republic
Molecular Systems Biology (MOSYS),
University of Vienna
Vienna, Austria
0000-0003-2918-9313

Vojtěch Bartoň
Department of Biomedical Engineering
Faculty of Eletrical Engineering and Communication,
Brno University of Technology
Brno, Czech Republic
RECETOX,
Masaryk University
Brno, Czech Republic
0000-0002-1918-2989

Dirk Walther
Max-Planck-Institute of Molecular Plant Physiology,
Potsdam, Germany
0000-0002-5755-9265

Wolfram Weckwerth
Molecular Systems Biology (MOSYS),
Vienna Metabolomics Center (VIME),
University of Vienna
Vienna, Austria
0000-0002-9719-6358

***Abstract* — Polyamines are known to be functionally involved in plant responses to stress conditions. Despite a large amount of physiological and genetic data, the mode of action of polyamines, especially putrescine, at the molecular level still remains unclear. An increasing number of studies point to a role of putrescine in stress defense of plants due to influencing the formation of reactive oxygen species (ROS) in cellular stress conditions. Moreover, putrescine has been found to modulate abscisic acid (ABA) biosynthesis at the transcriptional level in response to low temperature, revealing a novel mode of action of polyamines as regulators of hormone biosynthesis. This study presents a new holistic approach towards the analysis of putrescine metabolism using genome-wide association studies and the calculation of genetic risk scores extended by Boolean analysis in metabolites network as well as data mining from available databases and literature. It can lead to a better understanding of biological processes involved in adaptation of plants to environmental changes.**

Keywords — Diversification, GWAS, Genetic Risk Score, The 1001 Genomes Project, Reactive Oxygen Species

I. Introduction

Putrescine belongs to polyamines known to accumulate in response to stress and are known as small aliphatic amines, which can be found in all cells [1]. Putrescine is accumulated in response to stress and convincing evidence suggests a protective role for polyamines in cell immunity systems. A recent study by Liu and Alcázarby states that the polyamine putrescine (Put) accumulates during effector-triggered immunity [1].

A detailed integrated analysis of putrescine under cold stress was characterized in the study by Cuevas et. al. [2]. This study described in more detail the role of polyamines in response to low temperature. The study revealed that putrescine is required for plant survival under such limiting conditions by contributing to freezing tolerance.

However, the signalling pathways associated with polyamine functions are not yet fully revealed, described, and understood. Much needed is a proper description of the available data, combined with a computational analysis [3] – which is now possible even on huge volumes of data thanks to ever-better computing algorithms and resources. Data mining, combined with mathematical analysis can lead to an improved understanding of the regulators of hormone biosynthesis and the functions of putrescine in defence processes, and thus, to an improved plant breeding in general.

Hence, we first focus on the information contained in genomic DNA quantified as a prediction score called genetic risk score, which has been calculated for putrescine concentration obtained under a low-temperature condition of 6°C. Genetic risk score (GRS) [4] is an estimate of the cumulative contribution of genetic factors to a specific outcome of interest. This score takes into account the reported effect sizes of alleles. It is normalised by adjusting for the total number of risk alleles and effect sizes assessed. On the other hand, regarding genomic information, there is another approach of genomic analysis, which is now much more widespread in plant breeding – genome-wide association study (GWAS) [5, 6]. Due to the 1001 Genome Project [7], suitable information for this type of analysis such as GRS and GWAS is now available.

Our study focused on the analysis of the plant *Arabidopsis thaliana*. The study by Weiszmann et. al [8] describes the metabolic response to changing environmental factors (temperature) that differs significantly between ecotypes or cultivars of *Arabidopsis thaliana*. Moreover, we also follow up on the study [9], which showed significant differences in metabolite genome-wide association studies (mGWAS) from the study by Weiszmann et. al. [8] according to two temperature-related conditions. However, the present study focuses in more details on putrescine and extends it into a more comprehensive view via OMICs such as genomics and metabolomics analysis using GWAS [10], GRS [4] calculation, data mining in NCBI [11] and AraGWAS [12] in genomics, Boolean analysis and data mining in KEGG [13] in metabolomics.

II. Methods and Materials

The study deals with a comprehensive view of putrescine metabolism. Therefore, the study compares different approaches such as GWAS and GRS with known and available information from different databases. This comprehensive view was used to investigate the relationship between genetic factors and putrescine in *A. thaliana* to uncover links across different OMICs levels. An overview of the methodology is shown in the followed part of the 'Unification of known and explore information' section.

A. Plant materials

The study analyses data obtained for *A. thaliana* cultivated under different temperature-related conditions and are taken from the study by Weiszmann et. al. [8]. In addition, a dataset representing the natural variation of growth rates of *A. thaliana* and which was monitored together with dynamics of primary metabolites under moderate (16°C represented wild-type (WT)) and low (6°C) temperatures [9] was used, which we are extending with a more detailed focus on putrescine metabolism.

B. Genome-wide association study

As a follow-up to [8, 9], the basic GWAS function from package R/ rrBLUP [15] was used for the calculation of putrescine-GWAS. The main core is based on calculating maximum-likelihood solutions for mixed models of the form:

$$y = X\beta + Zg + S\tau + \varepsilon, \qquad (1)$$

where X, Z, and S present incidence matrices. X is of $n \times m$ size with unphased genotypes for n lines and m biallelic markers, coded as {-1,0,1}. β is a vector of fixed effects that model environmental factors. The variable g presents the genetic background of each line as a random effect, τ is the additive significant single nucleotide polymorphisms (SNPs) effect, ε is the residual error. [15]

C. Genetic risk score

GRS [4] is the most common approach to evaluate the cumulative effect of many genetic factors with small effects. It can estimate the probability (or risk) for the manifestation of an outcome of interest based on genetic variants. GRS-calculation to choose a number k of independent genetic variants with a strong association as risk predictors is a straightforward method [4]:

$$GRS = \sum_{i=1}^{k} \beta_i N_i, \qquad (2)$$

where β_i is from logistic regression analysis with additive genetic effect, multiplied by the number of risk alleles and N_i is for each locus.

The prediction accuracy of GRS is most often assessed by measuring the area under the receiver operating characteristic (ROC) curve (AUC) [4] but there are AUC limitations to consider. The potential accuracy [4] is known if all genetic factors are known, i.e.: (1) determination by the trait heritability and the correlation between the GRS and true genetic risk, and (2) determination by the quantity of genetic data and selection of genetic variants for prediction.

To calculate the GRS, a R script was implemented using R/PredictABEL [16] package. This was modelled by logistic regression and the prediction scores of the models were calculated as a function of each allele across all 241 natural accessions.

D. Boolean analysis in metabolomics

Boolean analysis such as Boolean network modelling can be extremely useful for the analysis an amount of qualitative data at the molecular level about individual components and revealing their possible unknown interactions [18, 19]. Nowadays, this understanding of experimental data is heavily involved in challenge regarding to the big data mining.

This study unified the biosynthesis of putrescine and the polyamines pathway taken from [17], expanded by the contribution of putrescine to defense in *A. thaliana* from [1]. It leads to simplifying and verifying the correctness of the pathway using the KEGG [13] database (pathway – map00330), see Fig. 1.

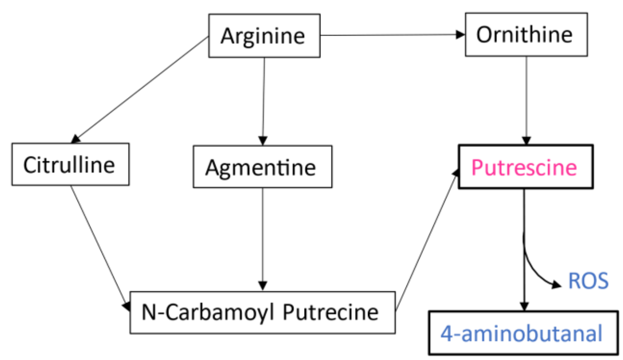


Fig. 1. Biosynthesis of putrescine and the polyamines taken from [17] expanded by new knowledge (highlight by blue) about the contribution of putrescine to defense in *A. thaliana* from [1].

The network shown in Fig. 1 was analysed using the Aeon tool [20]. The following two different situations were modelled. The first situation does not assume a negative effect of ROS in the form of feedback loop that can influence system stability. The second situation takes into consideration the possibility of the input value can be changed from logical 1 (true) to logical 0 (false) due to ROS-related feedback.

E. Unification of known and explore information

Intraspecific genetic variation is already able to explain a partial proportion of the metabolite-related phenotypes [21]. However, it is important not to underestimate the complex view as a whole of different processes across OMICs so called panomics [21]. Therefore, this study brings the unification of knowledge and computational analysis in genomics with available information and Boolean network modelling in metabolomics. The whole concept is shown in Fig. 2.

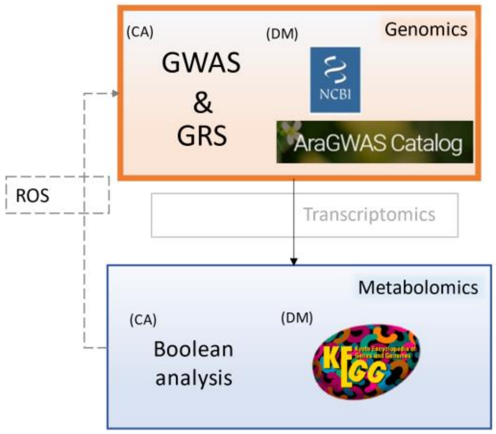


Fig. 2. A comprehensive methodology for metabolite analysis of putrescine linked to genomics information incorporated into the scheme of the central molecular dogma included steps from analysis such as computational analysis (CA) and data mining (DM).

Fig. 2 is based on basic knowledge of the central dogma of molecular biology, which is shown as a combination of the computational analysis and data mining process for a comprehensive analysis of putrescine. The diversification influenced DNA and radical oxygen species (ROS) – is known that putrescine elicits ROS-dependent activation [14] – have the potential for feedback to gene expression and genotype. Thus, influence the input parameters for metabolomic processes (simulated by us using simplified Boolean analysis).

III. Results and Discussion

A. *Genomics information for putrescine process in A. thaliana*

Firstly, GWAS related to metabolism of putrescine was performed (Fig. 3 and TABLE I. Fig. 3 includes results from computational analysis, which is expended by available data mining from NCBI and AraGWAS summarized in TABLE I.

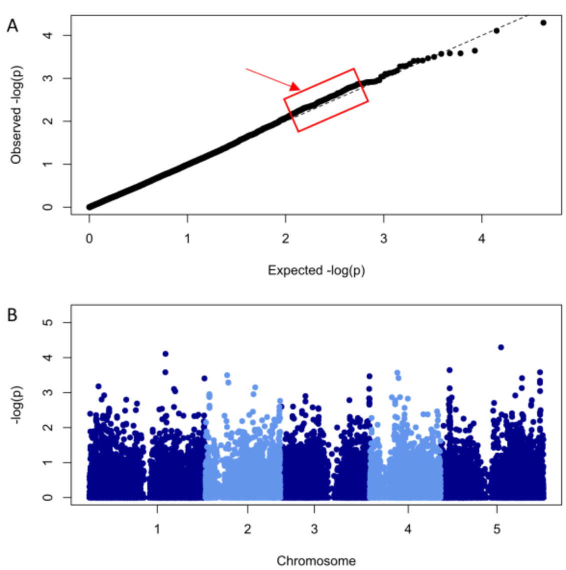


Fig. 3. (A) shows BLUP model constructed using rrBLUP package (B) is Manhattan plot based on the mixed model, with a significant change on the first and fifth chromosome. Each dot represents a SNP where light blue is the even number of a chromosome, dark blue is ann odd number of chromosomes.

Fig. 3(A) includes a red square highlighting the region of candidate for significant SNPs. However, in Fig. 3(B) is shown that the significant significance was not detected, so the study focused on the five best-ranked SNP by GWAS scores.

TABLE I includes the six the highest value of GWAS score. Surprisingly, the highest GWAS score relates to the SNPs that are located in an untranslated region (UTR 3'). Therefore, we can assume that a similar mode of regulation as described in current studies of polyamine function in mammalian cells [22] may also occur in the plant kingdom.

TABLE I. The six SNPs with the highest GWAS scores connected to putrescine

ID gene	GWAS Score	Chromosome	Data mining in AraGWAS
AT5G39560	4.29	5	UTR 3'
AT1G53520	4.11	1	Intron
AT5G07410	3.64	5	UTR 5'
AT5G65210	3.58	5	UTR 3'
AT1G53420	3.58	1	Synonymous coding
AT4G12030	3.57	4	UTR 3'

Regarding GRS, no significant predictive ability of a single allele was found, too. Total predictive value for all alleles in 6°C was calculated by 0.55 and for all alleles in 16°C was calculated by 0.45. The GRS-based predictive ability for putrescine at temperature-related 6°C was almost 10% higher than the GRS-based predictive ability at temperature-related 16°C. It can reflect easier prediction due to changes under cold stress conditions related to putrescine [2].

B. *Biosynthesis of putrescine given boolean network analysis*

Boolean analysis has shown the importance of feedbacks loops that can be across different level of OMICs (such as from metabolomics to genomics levels) due to ROS. Overall, we modelled two cases using Boolean analysis that were exported into SBML standardized format.

Firstly, Fig. 4 (A) shows the example of a simplified model without incorporation of ROS knowledge. In this case, there is possible 35 behaviour classes representing various possibilities of attractors. Secondly, Fig. 4 (B) shows the example of a simplified model which already incorporated ROS as single unit which can negatively affect genomic information. In second case, there is reduction of possible stable states represented by attractors into 15 behaviour classes representing various possibilities of attractors. By adding more complex information, we narrowed the stable state space of our simplified model focusing on metabolism of putrescine.

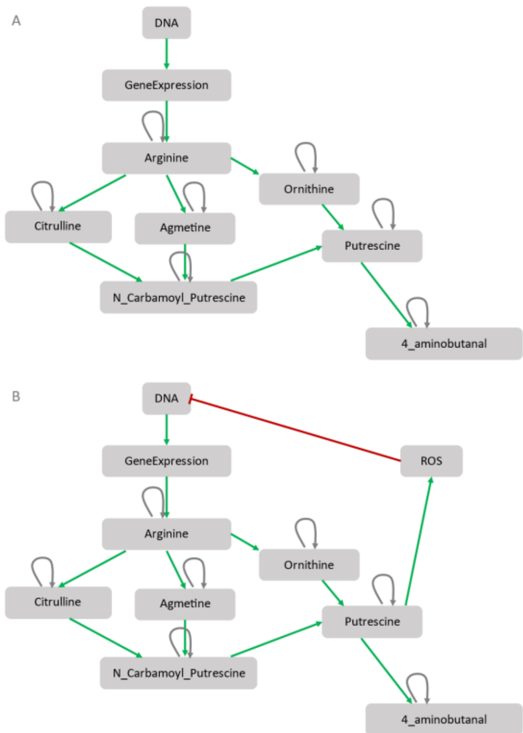


Fig. 4. Boolean networks models two different situation using AEON tool. (A) represents simplified boolean models without ROS feedback loop; (B) represents simplified boolean models with ROS feedback loop. The gray boxes represents nodes as variables in Boolean analysis. Edges represent regulations which can be positive (green arrow), negative (red edge) or unspecified (grey arrow).

C. *Holistic perspective of putrescine analysis*

The unification of all partial findings, see Fig. 5, brings the results of a comprehensive study that can better describe the real world. This study focused on the selected SNPs, which has scored as the highest in the putrescine-GWAS.

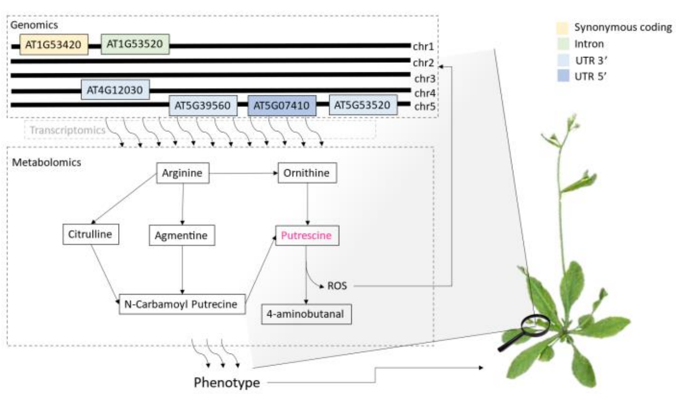


Fig. 5. Panomics insight into comprehensive putrescine analysis which combines genomics and metabolomics information.

These SNPs are involved in UTR regions such as in position 15,842,106 on the fifth chromosome or AT4G12030 codes protein required for the biosynthesis of methionine-derived glucosinolates and involved in the transport of 2-keto acids between chloroplasts and the cytosol [12]. Therefore, it seems appropriate to incorporate transcriptomic information as part of the continuation study as the next steps and inspired by polyamine regulation in mammalian cells [22].

IV. Conclusion

Combining biological knowledge with predictive modelling is considered a future cornerstone of systems biology, because predictive algorithms can offer detailed mathematical description of mechanisms that involve hidden or unknown interactions. However, before coming up with an understanding of interactions in complex systems resulting from emergent properties, it is necessary to understand the main influential factors that affect the properties of organisms.

Genetic variation using different algorithms such as GWAS and GRS can help us uncover important mechanisms in cellular defense under different stress conditions in plants. The study presents a powerful approach to reconnect this different panomics field back to its underlying genetics prediction based on metabolites such as putrescine. We combine information from genomics and metabolomics pathways using Boolean network analysis across a component such as ROS, which can influence processes across OMICs and thus change the overall inputs to metabolism.

Using a comprehensive view, we uncovered SNPs that are associated with the UTR region. Since this is an important area of gene expression regulation by alternative UTR regions, we can assume that adding transcriptomic information appears to be an important next step in analysing putrescine. Thanks to this, subsequently uncovering its role in signalling pathways and the overall effect on cold stress conditions.

Acknowledgment

This work has been supported by grant project FEKT/FIT-J-23-8274.

Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures.

References

[1] Liu, Changxin, and Rubén Alcázar. "A new insight into the contribution of putrescine to defense in Arabidopsis thaliana." *Plant Signaling & Behavior* 16, no. 4 (2021): 1885187.

[2] Cuevas, J.C., López-Cobollo, R., Alcázar, R., Zarza, X., Koncz, C., Altabella, T., Salinas, J., Tiburcio, A.F. and Ferrando, A., 2009. Putrescine as a signal to modulate the indispensable ABA increase under cold stress. *Plant signaling & behavior*, *4*(3), pp.219-220.

[3] Sidak, D., Schwarzerová, J., Weckwerth, W. and Waldherr, S., 2022. Interpretable machine learning methods for predictions in systems biology from omics data. *Frontiers in Molecular Biosciences*, *9*.

[4] Igo Jr, R.P., Kinzy, T.G. and Cooke Bailey, J.N., 2019. Genetic risk scores. *Current protocols in human genetics*, *104*(1), p.e95.

[5] Muqaddasi, Q.H., Brassac, J., Ebmeyer, E., Kollers, S., Korzun, V., Argillier, O., Stiewe, G., Plieske, J., Ganal, M.W. and Röder, M.S., 2020. Prospects of GWAS and predictive breeding for European winter wheat's grain protein content, grain starch content, and grain hardness. *Scientific reports*, *10*(1), p.12541.

[6] Ledesma-Ramírez, L., Solís-Moya, E., Iturriaga, G., Sehgal, D., Reyes-Valdes, M.H., Montero-Tavera, V., Sansaloni, C.P., Burgueño, J., Ortiz,

C., Aguirre-Mancilla, C.L. and Ramírez-Pimentel, J.G., 2019. GWAS to identify genetic loci for resistance to yellow rust in wheat pre-breeding lines derived from diverse exotic crosses. *Frontiers in Plant Science*, *10*, p.1390.

[7] Weigel, D. and Mott, R., 2009. The 1001 genomes project for Arabidopsis thaliana. *Genome biology*, *10*(5), pp.1-5.

[8] Weiszmann, J., Clauw, P., Jagoda, J., Reichardt-Gomez, I., Koemeda, S., Jez, J., Nordborg, M., Walther, D., Nägele, T. and Weckwerth, W., 2020. Plasticity of the primary metabolome in 241 cold grown Arabidopsis thaliana accessions and its relation to natural habitat temperature. *bioRxiv*, pp.2020-09.

[9] SCHWARZEROVÁ, J. Metabolite Genome-wide Association Studies of Arabidopsis Thaliana. In *Proceedings of the 27th Conference STUDENT EEICT 2021.* 2021. pp.41-44.

[10] Cantor, R.M., Lange, K. and Sinsheimer, J.S., 2010. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *The American Journal of Human Genetics*, *86*(1), pp.6-22.

[11] Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M. and Sirotkin, K., 2001. dbSNP: the NCBI database of genetic variation. *Nucleic acids research*, *29*(1), pp.308-311.

[12] Togninalli, M., Seren, Ü., Meng, D., Fitz, J., Nordborg, M., Weigel, D., Borgwardt, K., Korte, A. and Grimm, D.G., 2018. The AraGWAS Catalog: a curated and standardized Arabidopsis thaliana GWAS catalog. *Nucleic acids research*, *46*(D1), pp.D1150-D1156.

[13] Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. and Yamanishi, Y., 2007. KEGG for linking genomes to life and the environment. *Nucleic acids research*, *36*(suppl_1), pp.D480-D484.

[14] Liu, C., Atanasov, K.E., Arafaty, N., Murillo, E., Tiburcio, A.F., Zeier, J. and Alcázar, R., 2020. Putrescine elicits ROS‐dependent activation of the salicylic acid pathway in Arabidopsis thaliana. *Plant, Cell & Environment*, *43*(11), pp.2755-2768.

[15] Endelman, J.B., 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. *The plant genome*, *4*(3).

[16] Kundu, S., Aulchenko, Y.S., van Duijn, C.M. and Janssens, A.C.J., 2011. PredictABEL: an R package for the assessment of risk prediction models. *European journal of epidemiology*, *26*, pp.261-264.

[17] Smith, T.A., 1985. Polyamines. *Annual Review of plant physiology*, *36*(1), pp.117-143.

[18] Wang, R.S., Saadatpour, A. and Albert, R., 2012. Boolean modeling in systems biology: an overview of methodology and applications. *Physical biology*, *9*(5), p.055001.

[19] Cambiaghi, A., Ferrario, M. and Masseroli, M., 2017. Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration. *Briefings in bioinformatics*, *18*(3), pp.498-510.

[20] Beneš, N., Brim, L., Kadlecaj, J., Pastva, S. and Šafránek, D., 2020. AEON: attractor bifurcation analysis of parametrised boolean networks. In *Computer Aided Verification: 32nd International Conference, CAV 2020, Los Angeles, CA, USA, July 21–24, 2020, Proceedings, Part I 32* (pp. 569-581). Springer International Publishing.

[21] Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P. and Varshney, R.K., 2020. PANOMICS meets germplasm. *Plant Biotechnology Journal*, *18*(7), pp.1507-1525.

[22] Zou, T., Rao, J.N., Liu, L., Xiao, L., Yu, T.X., Jiang, P., Gorospe, M. and Wang, J.Y., 2010. Polyamines regulate the stability of JunD mRNA by modulating the competitive binding of its 3′ untranslated region to HuR and AUF1. *Molecular and Cellular Biology*, *30*(21), pp.5021-5032.

Estimation of myocardial conduction velocity using a coronary sinus catheter

1st Richard Ředina
Department of Biomedical Engineering
Brno University of Technology
Brno, Czech Republic
International Clinic Research Centre
St. Anna's Hospital
Brno, Czech Republic
ORCID: 0000-0002-4901-0504

2nd Marina Ronzhina
Department of Biomedical Engineering
Brno University of Technology
Brno, Czech Republic
ORCID: 0000-0003-1366-8336

***Abstract*—This paper deals with the measurement of the conduction velocity of the depolarization wave through the atrial myocardium. With the increasing number of electrophysiological procedures performed on people suffering from cardiac arrhythmias, the number of relapses is gradually rising. A reliable tool to predict these relapses is still being sought. One parameter for prediction could be the conduction velocity of the depolarization wave, which could reflect the degree of myocardial damage.**

Data from pediatric patients were used to measure it. With knowledge of the electrode placement on the intracardiac catheter and detection of atrial activity in the measured leads, the average myocardial conduction velocity along the catheter was calculated (v = 143.77 ± 21.08 cm/s).

This method brings simplicity and time saving to the whole process of measuring line speed. The results of this method should also be tested on an adult population where relevant models exist for comparison.

***Index Terms*—Intracardiac electrogram, Atria, Conduction velocity, Depolarization wave, Coronary sinus catheter**

I. Introduction

In recent years, there have been major developments in the field of electrophysiology. With the advent of more advanced diagnostics and the miniaturization of catheters used for atrial and ventricular ablation, the number of patients undergoing this treatment, which is a full-fledged alternative to long-established drug therapy, is increasing [1]. As the number of patients ablated increases, the number of arrhythmia recurrences due to patients previously undergoing the procedure also increases. Reliable prediction of patient relapses is one of the current challenges. To date, there is no 100% reliable parameter that directly relates to the likelihood of arrhythmia relapse. [2]

Myocardial conduction velocity could be one predictor of future relapses. The basic hypothesis is that in intact myocardium the conduction velocity of the depolarizing wave is higher than in myocardium affected by arrhythmias, ischemia or other degenerative influences [3]. The aim of this study is to develop and present a reliable method that will use the information commonly available from catheters to calculate conduction velocity in electrophysiological operations. Previously published work used whole atrial mapping or cellular models to calculate conduction velocity ([4], [5], [6], [7]). Despite higher accuracy, these methods are disproportionately more challenging. The method presented in this work creates a competitive model that allows real-time measurements during electrophysiological procedure.

II. Methods

A. Data

Records from twenty patients from the previously measured database [8] were selected for this study. The patients were aged 12-17 years. They underwent electrophysiological procedure at the Children's Hospital, University Hospital Brno, Brno, Czech Republic. The measurements were performed with the approval of the ethics committee. During the procedure, in addition to a 12-lead surface ECG, they had a 5-lead intracardiac ECG measured using a catheter inserted into the coronary sinus. The advantage of this database is that the atrial activity is annotated. Recordings, where only sinus rhythm was present, were used for this work. The recordings are therefore free of other arrhythmias or stimulation artifacts.

A 10-polar catheter was used for signal acquisition along with a St. Jude WorkMate 4.2 EP system with a sampling rate of 2000 Hz and a voltage resolution of 78 nV/LSB. The system incorporates a notch filter with a cutoff frequency of 50 Hz and an upper pass filter with a cutoff frequency of 0.1 Hz. The catheter used has electrodes arranged in a 2-8-2 mm configuration and the width of one electrode is 1 mm as can be seen in figure 1. As mentioned earlier, the catheter was placed in the coronary sinus whose orifice is located in the right atrium (see figure 2). This vein runs in the sulcus coronarius and runs around the perimeter of the left atrium. Due to this anatomical position, the depolarizing wave propagates along the catheter from its proximal to its distal end.

The electrodes placed on the catheter form five pairs from which bipolar leads are measured. These are designated in the records as cs1-2, cs3-4, cs5-6, cs7-8, cs9-10. The electrodes are numbered from the distal end of the catheter. The order of the leads and their spacing will be crucial for subsequent

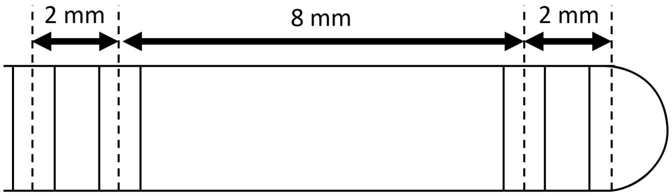


Fig. 1. Diagram of electrode placement on the intracardiac catheter with indicated spacing.

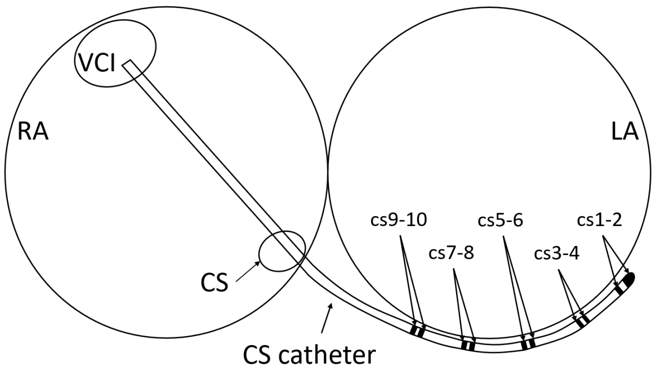


Fig. 2. Placement of coronary sinus catheter in the heart. Catheter comes via inferior vena cava (*VCI*), through right atrium (*RA*), it is inserted to another vein - coronary sinus (*CS*), which runs around the left atrium (*LA*). The location of the leads is shown on the catheter.

analysis and calculation of the conduction velocity of the depolarizing wave through the atrial myocardium.

The signal from the catheters measured during physiological sinus rhythm can be seen in the figure 3. The signal can be divided into three parts. The first is the isoelectric line, which is present for most of the cardiac cycle. It begins approximately with the end of the S wave in the surface ECG, continues during the T wave, and ends during the P wave. During the P wave, high-frequency activity can be found in the intracardiac leads during the surface P wave. This corresponds to the course of the depolarization wave. This is followed by a low-frequency wave corresponding to the depolarization of the ventricles. This wave may not be present in all intracardiac leads. Moreover, its frequency composition may overlap with atrial activity and therefore its filtering may be difficult.

B. Atrial activity Detection

The main aim of the study is to determine the velocity of the depolarization wave propagating along the atria along the catheter placed in the coronary sinus. The aim of the work with the measured signal will be mainly to correctly localize the peaks of high-frequency activity in the five leads measured by this catheter. The signal analysis is based on two important facts:

1) The depolarization wave appears in the leads gradually from the proximal end of the catheter to its distal part.
2) The character of the signal during ventricular activity has a frequency maximum below 50 Hz.

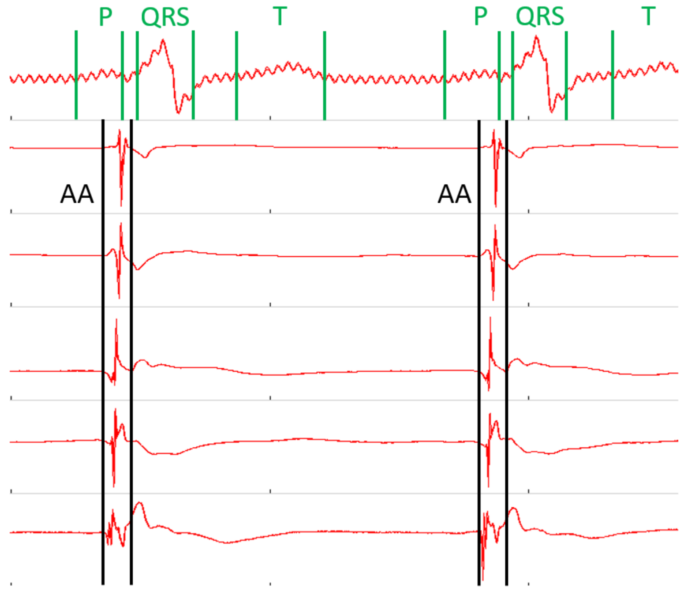


Fig. 3. Comparison of intracardiac signal vs surface ECG. The atrial activity (*AA*) in CS leads corresponds with the second half of P wave in surface ECG. The image shows a cutout 1.25 s long. Leads from the top: V6, cs1-2, cs3-4.cs5-6, cs7-8, cs9-10. Depolarization propagates from the lead cs9-10 to cs1-2. The signal was plotted using SignalPlant v.1.2.7.9 [9].

The first step in the processing of intracardiac signals is filtering in order to highlight the atrial component and to suppress as much as possible the ventricular activity or possible noise caused most often by electric network noise. For filtering, a wavelet transform was chosen to divide the signal spectrum into a predefined number of octave bands [10].

To suppress the low frequency components, it is sufficient to obtain the first band in which the frequencies from $f_s/4$ to $f_s/2$ are captured. This step reliably removes from the signal most of the chamber activity that may be present in the signals. However, depending on how far in the coronary sinus the catheter is inserted, it is possible that the distal end is already too close to the left ventricle and therefore high frequency ventricular activity may be present in the signal. Therefore, after filtering, two peaks may occur in the signal close together. Their resolution will be subsequently addressed.

The next step was to create the envelope of the filtered signal. Using the Hilbert transform, an analytical signal was obtained whose magnitude is the envelope of the signal [11]. This was then smoothed with a low-pass filter to remove as many local outliers as possible. In this way, an envelope was created for each of the five signals. The result of this step is shown in the figure 4.

From the generated envelope, it is already possible to detect local maxima in each signal. Since the depolarizing wave propagates along the catheter from its proximal end, it is possible to start with peak detection in lead cs9-10 and adjust the detection in subsequent leads. Two conditions have been chosen that the peak must satisfy to be detected:

1) It must be greater than one-sixth of the maximum value taken by the signal envelope.

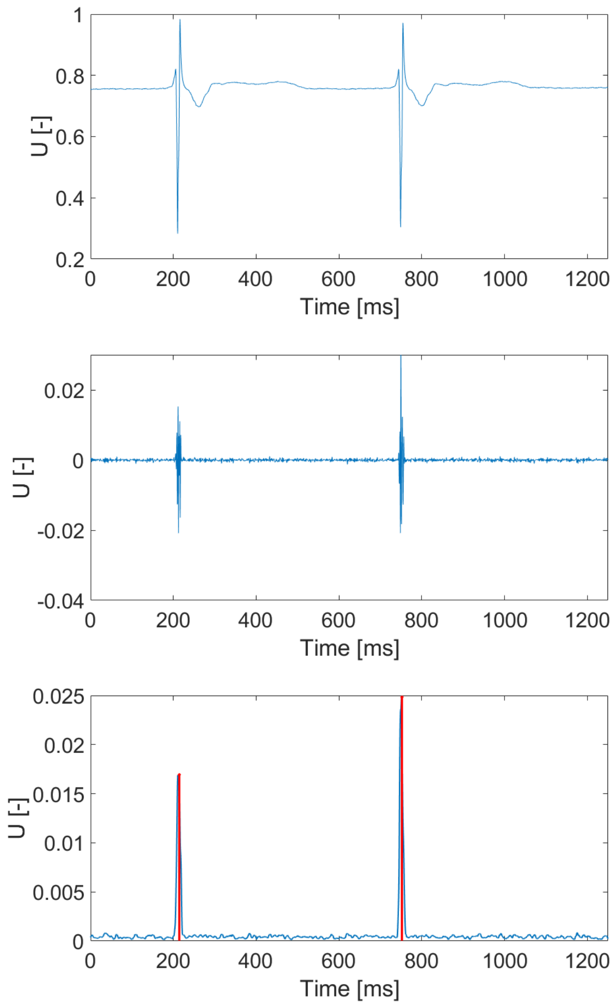


Fig. 4. Intracardiac signal procedure. From top to bottom: a) scaling the voltage values of the raw signal to an interval from 0 to 1, b) filtering using a wavelet transform in order to highlight the high-frequency component of the signal, c) creation of an envelope of filtered signal (blue) from which the peaks (red lines) were detected.

2) Two adjacent peaks must be at least 0.45 s apart.

The first condition is based on empirical adjusting of the detection threshold. The second condition is based on the fact that the heart rate in children during sinus rhythm may be higher than in adults, but still should not exceed 130 bpm.

From the set of detected positions, those whose distance from two adjacent positions was less than one and a half times the standard deviation of the differences subtracted from the mean value of the differences were removed. This reduced the number of false positive detections that could bias the final conduction velocity calculation.

In subsequent leads, the detection of envelope peaks was modified so that in each subsequent lead the peak was searched for in a local neighborhood defined by the position of the previous peak. The neighborhood was defined from -15 to +35 milliseconds from the previous peak. Again, a local envelope maximum was found in this region. The neighborhood also considers the short signal segment before the preceding peak for the reason that cardiac tissue is not an isotropic environment for the propagation of the depolarizing wave and thus deviations from the expected propagation direction may occasionally occur.

C. Velocity Estimation

The position of atrial activity across the five leads in the intracardiac ECG detected in the previous section, together with the knowledge of the electrode placement on the catheter, is used for the final calculation of the conduction velocity of the depolarizing wave through the myocardium. Despite the possible presence of anatomical variations, we will assume that the catheter runs around the surface of the atria, with all electrodes insistent on the myocardium. Because of this, we can consider the distances of the leads on the catheter to correspond to the distance between the measured sites on the atrial wall.

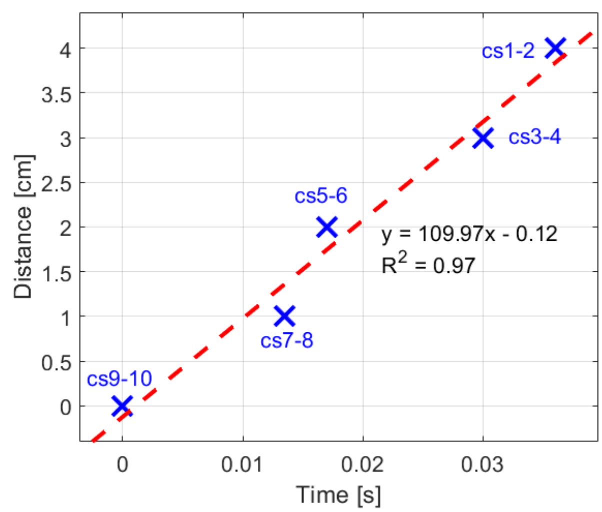


Fig. 5. Regression model for one cardiac cycle. The individual points represent the position of the maximum in each of the five leads. The leads are spaced one centimeter apart. The slope of the obtained line corresponds to the propagation speed of the depolarization wave for this case which is approximately 110 cm/s. The high value of the R^2 parameter indicates that the linear regression model explains most of the variability captured in the data and is therefore appropriate for this case.

One position obtained from each lead was always used to calculate the velocity. The position from catheter cs9-10 was marked as point [0,0]. The other leads were followed one centimeter at the lead. From the five points plotted on the distance-time graph, the coefficients of the line were obtained by linear regression. The slope of this line corresponds to the propagation velocity of the depolarizing wave in cm/s. This procedure was repeated for each cardiac cycle within a single signal. The resulting velocity for each patient is given by the average of the all velocities. An example for one cardiac cycle is shown in the figure 5.

III. Results

The outcome of the entire conduction velocity measurement depends on the accuracy of the detection of atrial activity in the intracardiac recordings. This was assessed using standard parameters such as sensitivity (true positive rate, TPR), positive predictive value (PPV) or F1 score [12]. The calculation of these parameters is summarized in equations (1), (2) and (3). As a ground truth detection, the position annotations from the database were taken.

$$TPR = \frac{TP}{TP + FN}, \tag{1}$$

$$PPV = \frac{TP}{TP + FP}, \tag{2}$$

$$F1score = \frac{2 * TP}{2 * TP + FN + FP}, \tag{3}$$

where true positive (TP) is the number of correctly detected atrial activity episodes, false positive (FP) is the number of detected atrial activity episodes not presented in the signal and false negative (FN) is the number of atrial activity episodes present in the signal, but not detected by the algorithm.

Table I summarizes these calculated parameters. As can be seen, the chosen procedure has a relatively high sensitivity and a 100% positive predictive value. Thus, almost all episodes present in the signals were detected correctly. The algorithm did not detect any false positive episodes. False negative detections could be further eliminated by adjusting the detection thresholds.

TABLE I
Accuracy parameters of peak detection in intracardiac leads

Attribute	Value
Number of true positive	228
Number of false positive	0
Number of false negative	21
True positive rate [%]	91.57
Positive predictive value [%]	100.00
F1 score [%]	95.60

For each of the twenty patients tested, the conduction velocity of the depolarizing wave through the myocardium was calculated. These results are summarized in table II. Because of the variability in heart rate, the number of cardiac cycles from which the velocity was calculated was also different for each patient. The table also captures other descriptive statistics such as the mean and the standard deviation of the velocities.

The calculated velocities have a normal distribution (143.77 ± 21.08 cm/s), as can be seen in figure 6, where the result of the Shapiro-Wilk test is mentioned. Along with this, the result of the student's one sample t-test against the velocity value found in the available literature (52.7 ± 25.3 cm/s [5], 91.0 ± 44 cm/s [13]) was significantly different. The

TABLE II
Descriptive statistics of the obtained results

Attribute	Value
Number of subjects	20
Mean number of heart cycles per patient	11.4
Minimal conduction velocity [cm/s]	95.26
Maximal conduction velocity [cm/s]	198.81
Mean conduction velocity [cm/s]	**143.77**
Minimal standard deviation of conduction velocity [cm/s]	3.17
Maximal standard deviation of conduction velocity [cm/s]	46.60
Mean standard deviation of conduction velocity [cm/s]	**21.08**

result of [14] showed that the conduction velocity can reach up to 150 cm/s. However, this work focused on the His bundle, not the atrial myocardium.

On the other hand, in papers [4] and [15] the conduction velocity in some patient groups reached up to 150 cm/s ± 50 cm/s. Such conduction velocities are already close enough to the values measured in the presented way (p = 0.366).

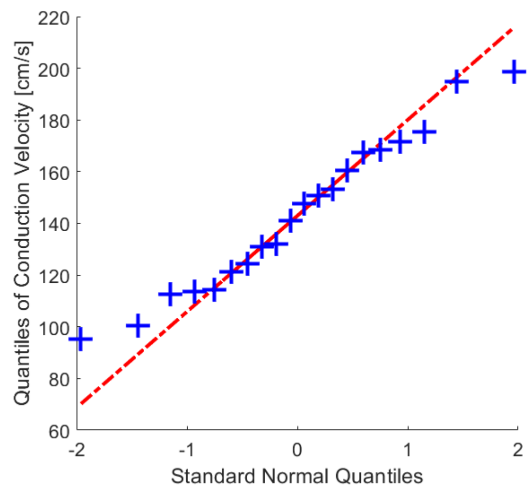


Fig. 6. Quantile-quantile plot of conduction velocities within the measured dataset. The data have a normal distribution, which was also verified using the Shapiro-Wilk's test (W = 0.96, p-value = 0.7).

IV. Conclusion and Discussion

In this paper, a new method for measuring the conduction velocity of the depolarizing wave through the atrial musculature is described. The main advantage of this method lies in the use of data measured with a catheter placed in the coronary sinus. The method is simple, fast and uses signals that are commonly measured. This catheter is commonly used in electrophysiological procedures aimed at removing foci of cardiac arrhythmias by applying radiofrequency energy. This work is also unique because of the data used, which comes from the child population. However, this fact is also a weakness of the work because to the best of our knowledge no similar study has been conducted in children, so the measured conduction velocity can only be compared to results on the adult population.

Compared to conduction velocity measured on hearts from adults, the rate measured by this method is significantly higher compared to some of the literature. This could be due to several factors:

1) Age of the measured patients. The heart, like all organs, is subject to more wear and tear with increasing age. Taking this fact into account, one may consider the possible higher myocardial conduction velocity in children compared to adults who already have a heart affected to some extent. This hypothesis is also supported by the fact that experiments on isolated hearts have shown that as the ischaemia-affected cardiac wall is progressively remodelled, the conduction velocity through the affected area also decreases. To clearly confirm or refute this hypothesis, a group of adult patients undergoing electrophysiological procedures would also need to be tested. This data is not currently available. If the hypothesis were verified, conduction velocity could be considered as a marker of myocardial damage.
2) Number of patients studied. The number of patients whose data was used for this work is limited by the size of the available database and the fact that it is primarily a database of arrhythmias. Pure sinus rhythm records are also present but are not the main domain. Thus, it is possible that the twenty patients selected do not adequately capture the variability in the entire pediatric population. An option to broaden the patient selection could be to select certain types of regular arrhythmias (such as atrioventricular nodal tachycardia, atrioventricular tachycardia, atrial flutter, etc.) and adjust the peak detection parameters for these specific cases. Another option would be to include recordings with paced rhythms. However, these recordings are problematic for two reasons. The first is that the stimulation is often performed with the catheter being used. While this pinpoints the exact moment in time when the wave is generated, it also creates an artifact that is reflected in other leads and can complicate detection in them. The second problem is that any electrode on the catheter can theoretically be used for stimulation. If electrodes other than the outermost electrodes are used for stimulation, the depolarizing wave would propagate to both ends of the catheter, reducing the accuracy of the resulting velocity measurement.

In the two papers mentioned, the conduction velocity was close to the measured value presented in this paper. This was mainly in the control groups of patients. The children whose recordings were used for this work could also be considered as otherwise healthy control subjects, because the main reason for performing the electrophysiological procedure was structural malformations, especially accessory pathways leading to WPW syndrome. If we focus on the atrial myocardium itself, it can be assumed that the structural damage to it was not very high given the young age.

The reason for the relatively high variability in velocities could be anisotropy of cardiac tissue. The cardiac musculature does not form a unit in which the conduction velocity is the same in each direction. It is therefore possible that the average propagation velocity will be lower, but equally there may be areas of muscle where conduction will be faster. This hypothesis could be tested by measuring with multipolar catheters that would map both atria and then calculate specific velocity values in different directions.

The age of the patients, number of patients and high variability caused by anisotropy of cardiac tissue are probably the biggest issues that would lead to a refinement or at least verification of the correctness of the method used. The hypothesis that conduction velocity could be used as a predictive factor to determine arrhythmia relapse in patients cannot be reliably confirmed. However, there is certainly room for further investigation of this idea.

Acknowledgment

Brno Ph.D. Talent Scholarship Holder – Funded by the Brno City Municipality.

References

[1] PALAMA, Zefferino, et al. Tailoring the ablative strategy for atrial fibrillation: a state-of-the-art review. Cardiology Research and Practice, 2022, 2022.

[2] LIŻEWSKA-SPRINGER, Aleksandra, et al. Echocardiographic predictors of atrial fibrillation recurrence after catheter ablation: a literature review. Cardiology Journal, 2020, 27.6: 848-856.

[3] KLEBER, Andre G., et al. Changes in conduction velocity during acute ischemia in ventricular myocardium of the isolated porcine heart. Circulation, 1986, 73.1: 189-198.

[4] MIYAMOTO, Koji, et al. Bipolar electrogram amplitudes in the left atrium are related to local conduction velocity in patients with atrial fibrillation. Europace, 2009, 11.12: 1597-1605.

[5] PADILLA, Jairo Rodriguez, et al. Orientation of conduction velocity vectors on cardiac mapping surfaces. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology: Journal of the Working Groups on Cardiac Pacing, Arrhythmias, and Cardiac Cellular Electrophysiology of the European Society of Cardiology, 2023, euac259-euac259.

[6] VIGMOND, E. J., et al. Solvers for the cardiac bidomain equations. Progress in biophysics and molecular biology, 2008, 96.1-3: 3-18.

[7] CANTWELL, Chris D., et al. Techniques for automated local activation time annotation and conduction velocity estimation in cardiac mapping. Computers in biology and medicine, 2015, 65: 229-242.

[8] REDINA, Richard, et al. Arrhythmia database with annotated intracardial atrial signals from pediatric patients undergoing catheter ablation. Age [years], 14: 12.0-17.0.

[9] PLESINGER, F., et al. SignalPlant: an open signal processing software platform. Physiological measurement, 2016, 37.7: N38.

[10] HEJČ, Jakub, et al. A wavelet-based ECG delineation method: adaptation to an experimental electrograms with manifested global ischemia. Cardiovascular engineering and technology, 2015, 6: 364-375.

[11] XU, Wansong; DU, FANYU. A robust qrs complex detection method based on shannon energy envelope and hilbert transform. Journal of Mechanics in Medicine and Biology, 2022, 22.03: 2240013.

[12] FAWCETT, Tom. An introduction to ROC analysis. Pattern recognition letters, 2006, 27.8: 861-874.

[13] PARK, Jae H., et al. Electrophysiologic characteristics of complex fractionated atrial electrograms in patients with atrial fibrillation. Journal of cardiovascular electrophysiology, 2009, 20.3: 266-272.

[14] KUPERSMITH, Joel; KRONGRAD, Ehud; WALDO, Albert L. Conduction intervals and conduction velocity in the human cardiac conduction system: studies during open-heart surgery. Circulation, 1973, 47.4: 776-785.

[15] SANDERS, Prashanthan, et al. Electrical remodeling of the atria in congestive heart failure: electrophysiological and electroanatomic mapping in humans. Circulation, 2003, 108.12: 1461-1468.

Chest X-ray Image Analysis using Convolutional Vision Transformer

Anzhelika Mezina
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
xmezin00@vut.cz

Radim Burget
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
burgetrm@vut.cz

***Abstract*—In recent years, computer techniques for clinical image analysis have been improved significantly, especially because of the pandemic situation. Most recent approaches are focused on the detection of viral pneumonia or COVID-19 diseases. However, there is less attention to common pulmonary diseases, such as fibrosis, infiltration and others. This paper introduces the neural network, which is aimed to detect 14 pulmonary diseases. This model is composed of two branches: global, which is the InceptionNetV3, and local, which consists of Inception modules and a modified Vision Transformer. Additionally, the Asymmetric Loss function was utilized to deal with the problem of multilabel classification. The proposed model has achieved an AUC of 0.8012 and an accuracy of 0.7429, which outperforms the well-known classification models.**

***Index Terms*—deep learning, multilabel classification, chest X-ray images, Vision transformer, InceptionNetV3**

I. INTRODUCTION

With a growing amount of information in the clinical field, it is getting more difficult to process it in time. Especially, with the beginning of the COVID-19 pandemic people used to utilize the health system much more frequently. In most cases, patients had health problems related to pulmonology, consequently, these departments of hospitals were mostly crowded.

One of the most common ways how to examine patients is to analyze their X-ray images. In recent decades, a lot of techniques for image analysis in the clinical field have been developed, including the application of deep learning (DL) methods [1]. Many researchers focused on the detection of pneumonia or COVID-19 and propose a lot of approaches using DL methods [2]. Here effectiveness of artificial intelligence (AI) was proved: it can find relations in a large amount of data and correctly detect disease. However, last 3 years there was not so much attention to other pulmonary diseases which still can appear, even as a consequence of COVID-19, for example, fibrosis or nodule.

This work focuses on the design of the neural network, which would detect 14 pulmonary diseases in X-ray images using a combination of InceptionNetV3 and a modified Vision Transformer by adding convolutional layers. For this aim, the large-scale dataset, ChestX-ray14 [3], was utilized.

To prove the effectiveness of our architecture, a comparison with well-known architectures for classification task is also provided.

The main contributions of this paper are:

- The architecture of a neural network for the detection of pulmonary diseases from X-ray images is proposed.
- The problem of multilabel classification is taken into consideration and the specific loss function is used for this task – Asymmetric Loss function.

The rest of this paper is structured as follows. Some recent works are introduced in Section II. The proposed architecture of the neural network is described in Section III. Section IV provides the results and discussion. Section V concludes the work.

II. RELATED WORKS

A. Pulmonary diseases classification using Deep learning

Nowadays, pulmonary disease detection and classification in X-ray images is an active area of research. Currently, most approaches are aimed to detect COVID-19 and pneumonia, however, some works focus on the detection of other pulmonary diseases. For example, work [4] introduced the architecture, based on DenseNet-169, which is aimed to classify X-rays into COVID-19, Pneumonia, Tuberculosis and normal. The achieved accuracy is 91%.

The well-known classification models were also trained in approach [5] for COVID-19 detection – EfficientNetB1, NasNetMobile and MobileNetV2, which achieved an accuracy of 96.1%, 94.81% and 93.96% respectively.

Another possible solution is the utilization of feature extraction. Approach [6] used ResNet50 as the backbone. Additionally, the authors applied a shuffle attention module, which is supposed to focus on discriminative features of abnormality. The achieved results are promising – 96% accuracy for COVID-19 detection.

Also, the work [7] uses pretrained EfficientNetB0, EfficientNetB1, and EfficientNetB2 as multichannel feature extraction. After that, the feature fusion module processes the extracted features and the final classification is provided by the stacking ensemble approach.

One of the recent works [8] introduced the modified MobileNetV2 for more accurate detection of pulmonary diseases.

AUC metric was used as the main performance measure, and the model achieved 0.923.

B. Transformers in medical image processing

In recent years, transformers are getting very popular in the research area. Initially, this architecture was applied for Natural Language Processing (NLP), but with the introduction of Vision Transformers (ViT) [9], transformers found a place in image processing, including in the medical field of research.

In this way, work [10] utilized and modified ViT for the detection of pulmonary diseases in chest X-ray images. The modification was inspired by the architecture of the ResNet model where some skip connections are used. The proposed architecture was aimed to classify images into COVID-19, Pneumonia, and normal.

Another approach [11] for the detection of COVID-19 is based on Data-Efficient Image Transformer. The model consists of three main blocks: an embedding layer, a Siamese encoder, and a decoder. The overall accuracy of the model is 94.62%.

The multilabel classification of chest X-ray images using a pyramid Vision Transformer was also considered in work [12]. The model can capture short and long-range visual information using self-attention. The achieved AUC score is 94.6% on the Catheter dataset.

The combination of ResNet18 and ViT is used in [13]. ResNet18 architecture extracts features which are encoded by Transformer. The final classification is performed with Multi-Layer Perceptron (MLP).

III. Methodology

A. Data preprocessing

For this experiment, the large well-known dataset, ChestX-ray14 [3], is used. The first step is cropping the image to the area of interest, which is the part with the lungs. It was performed using a pre-trained U-Net for lung segmentation. The next step is resizing to 384px $\times$ 384px. The total number of training images in the dataset is 36,024, and testing – 15,735. The validation set is 20% of the training set. The total number of classes is 14 and it is important to note, that one image can contain one or more diseases, which is why the dataset is prepared for a multilabel problem. The information about class distribution is presented in Table I. Finally, before feeding into the neural network, the image was augmented (horizontal flip, height and width shifting, rotation, and zoom).

B. Baseline

To evaluate the effectiveness of the proposed model, well-known architectures, such as DenseNet121 [14], EfficientNetB4 [15], ResNet101 [16], VGG16 [17] and InceptionNetV3 [18], are also trained. These models are used as feature extractors and the final classification is performed with three dense layers with 64, 32, and 14 units. The last layer has a sigmoid activation function. The used loss function is binary cross-entropy (BCE), and an optimizer is AdamW with a learning rate – 0.0001, and weight decay – 0.00001.

TABLE I
Distribution of classes.

Pathology	Train set	Test set	Total
Infiltration	13782	6112	19894
Effusion	8659	4658	13317
Atelectasis	8280	3279	11559
Nodule	4708	1623	6331
Mass	4034	1748	5782
Pneumothorax	2637	2665	5302
Consolidation	2852	1815	4667
Pleural Thickening	2242	1143	3385
Cardiomegaly	1707	1069	2776
Emphysema	1423	1093	2516
Edema	1378	925	2303
Fibrosis	1251	435	1686
Pneumonia	876	555	1431
Hernia	141	86	227

C. Proposed architecture

The proposed model is aimed to deal with multilabel classification with 14 labels, taking into consideration the imbalance of the dataset. The model consists of two branches: global and local. The general scheme of the proposed model is introduced in Figure 1.

The global branch is represented by InceptionNetV3. Because of its structure, it is possible to extract features of different sizes. After feature extraction by InceptionNetV3, the Global Max Pooling layer and Dense layer process the extracted information.

The local branch consists of Inception modules, modified ViT and Global Average Pooling. Three inception modules extract the local information from an input image and reduce the dimension of extracted feature maps. The next step is performed with a modified Vision Transformer. Firstly, the extracted features are encoded. After that, 8 Transformer layers process them. Figure 2 shows detail of a local branch, including the Transformer layer. This layer is composed of Layer Normalization, Convolutional, Multihead Attention and Add layers. The main difference from the original ViT transformer layer is added Convolutional layers after Layer Normalization. The last part of Transformer layers is MLP. This one consists of Dense layers with 64 and 32 units and Dropout layers with a drop rate of 0.2. After that, the output is processed by Layer Normalization, Global Average Pooling and Dropout and following MLP, which also has Dense layers 64 and 32 units and Dropout layers with a drop rate 0.2.

The final part of the proposed model is the concatenation branch composed of Dense and Dropout layers. The last layer is dense with 14 units and a sigmoid activation function [19], which can produce results in the range of 0 to 1.

As it was mentioned, the solved problem is multilabel classification. Because of that, the specific loss function was used – Asymmetric Loss (ASL) [20]. Mathematically, ASL

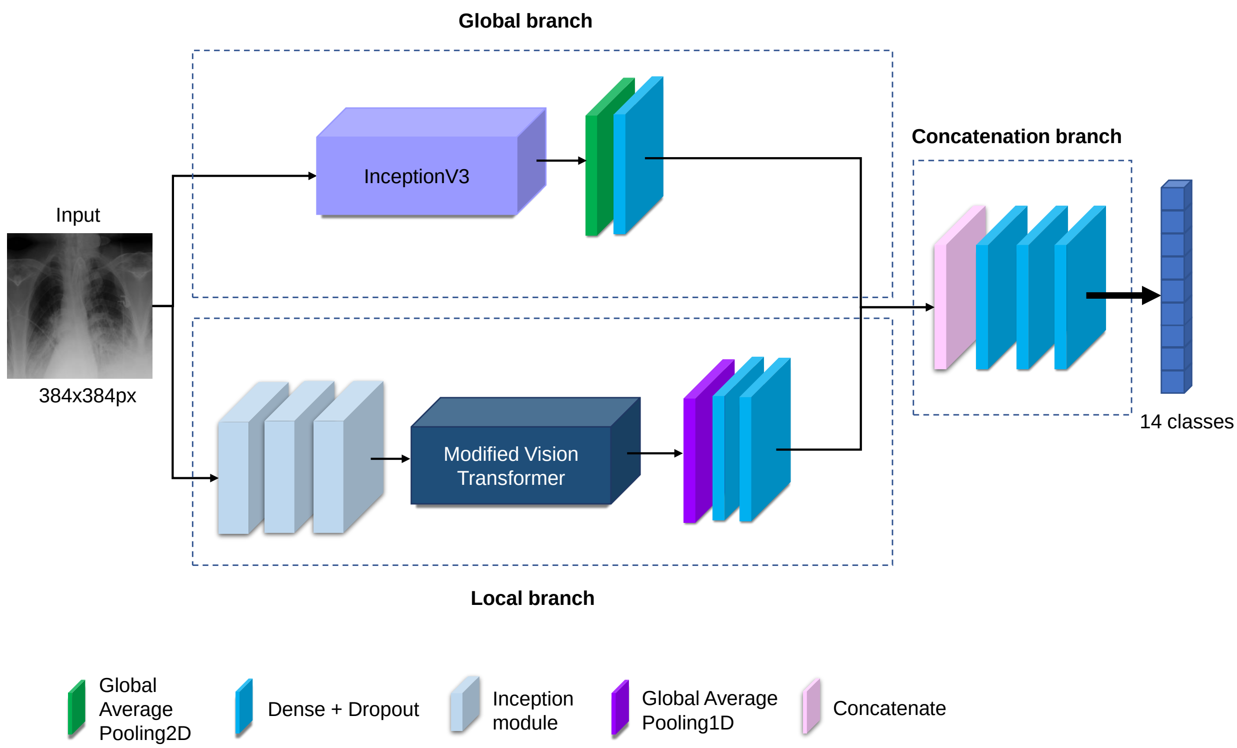


Fig. 1. Architecture of proposed model.

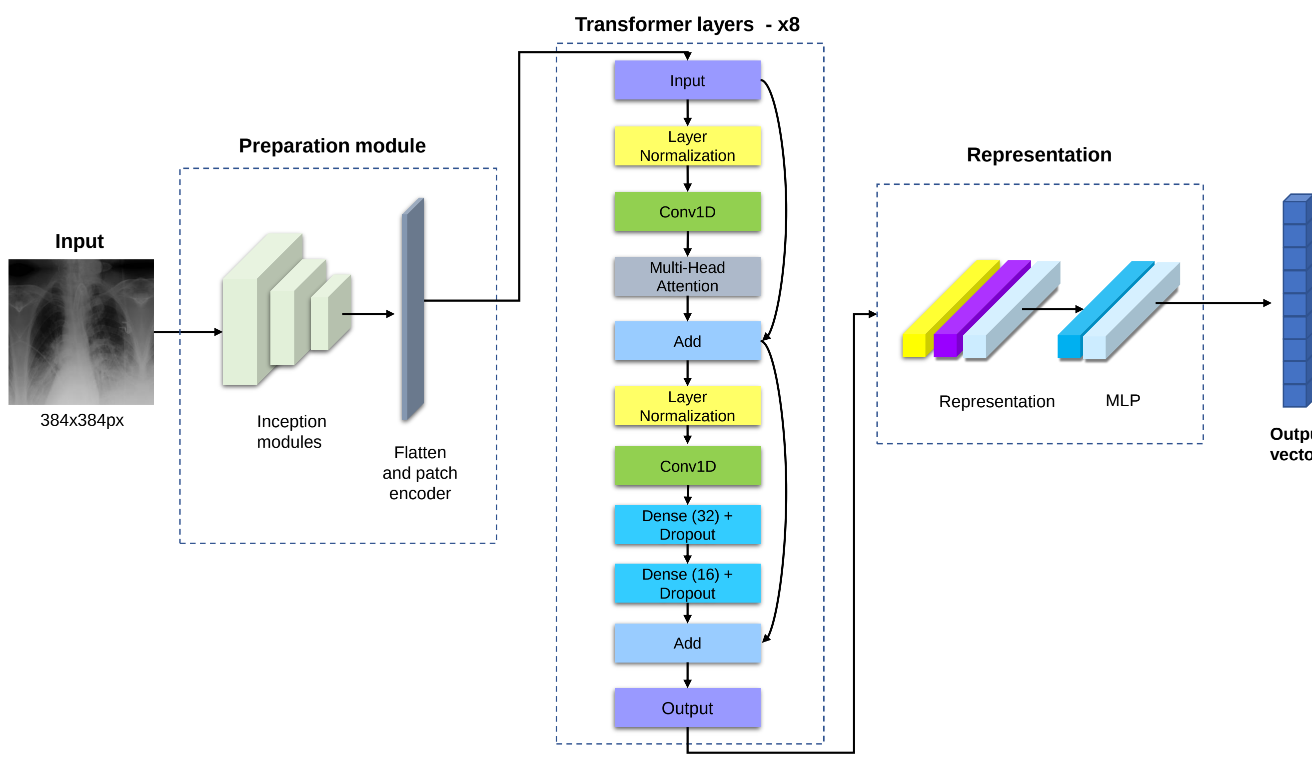


Fig. 2. Architecture of modified Vision Transformer.

can be described as:

$$ASL = \begin{cases} L_+ = (1-p)^{\gamma_+} \log(p), \\ L_- = (p_m)^{\gamma_-} \log(1-p_m), \end{cases} \tag{1}$$

where L_+ and L_- are the positive and negative loss parts, p is an output probability from the network, p_m is a shifted probability, γ is a focusing parameter. The set up parameters are γ_- is 5, γ_+ is 1, clip is 0.001. The hyperparameters are: number of epochs – 50, steps per epoch – 1000, an optimizer is AdamW with a learning rate – 0.0001, and weight decay – 0.00001.

IV. Results and evaluation

A. Metrics

To evaluate the performance, some metrics are used: accuracy, sensitivity, specificity, precision, and F1. Their definitions are defined as:

$$Accuracy = \frac{TN+TP}{TP+TN+FP+FN}, \tag{2}$$

$$Precision = \frac{TP}{TP+FP}, \tag{3}$$

$$Sensitivity = \frac{TP}{TP+FN}, \tag{4}$$

$$Specificity = \frac{TN}{TN+FP}, \tag{5}$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}, \tag{6}$$

where TN – True Negatives, TP – True Positives, FP – False Positives, FN – False Negatives.

B. Results and discussion

Table II summarizes the achieved results by the proposed model and models used for baseline. It can be seen, that our model outperforms other architectures with the following achieved results: AUC – 0.8012, specificity – 0.7452, accuracy – 0.7429, precision – 0.2627, F1 – 0.3554.

The models in baseline perform well for different metrics: VGG16 is better for AUC (0.7797), precision (0.2433), F1 (0.3357), and Sensitivity (0.7427), EfficientNetB4 is the best according to specificity (0.7221), accuracy (0.7223).

More detailed results with AUC metric are introduced in Table III and Figure 3. Compared to other models, the proposed neural network has the best results for almost all diseases: Atelectasis – 0.7635, Cardiomegaly – 0.8825, Consolidation – 0.6907, Edema – 0.8310, Effusion – 0.8114, Emphysema – 0.8988, Hernia – 0.9065, Infiltration – 0.7054, Mass – 0.8098, Nodule – 0.7692, Pleural Thickening – 0.7666, Pneumonia – 0.6926, Pneumothorax – 0.8671. EfficientNetB4 has achieved better results for Fibrosis – 0.8263.

Here can be concluded, that the most problematic diseases for detection are: Consolidation, Infiltration, and Pneumonia. On the other hand, Hernia, Cardiomegaly, Emphysema and Pneumothorax are easily detected and classified by a neural network.

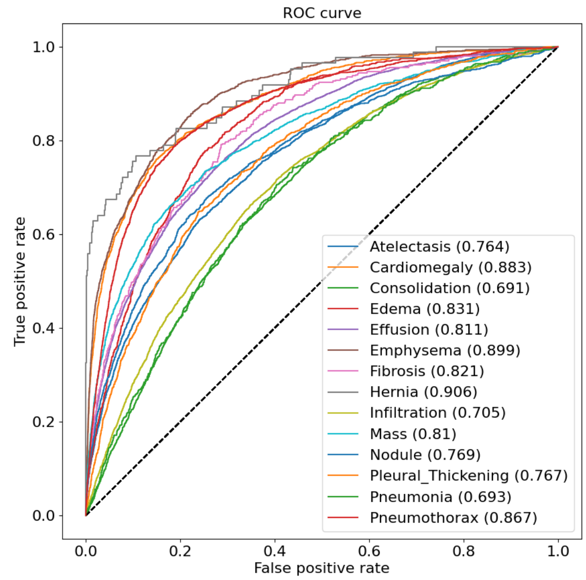


Fig. 3. ROC-AUC for the proposed neural network.

Also, it is worth mentioning, that it is possible to compare the results of pure InceptionNetV3 and a combination of this model with modified ViT using ASL. Here it is obvious, that the results are more accurate: AUC is better for 0.05, sensitivity – 0.02, specificity – 0.06, accuracy – 0.06, precision – 0.03, F1 – 0.04. Table III also proved the effectiveness of the proposed model compared with InceptionNetV3.

V. Conclusion

This work introduced the model of the neural network, which detects 14 pulmonary diseases. This model captures global and local features from input X-ray images using InceptionNetV3 and a modified Vision Transformer model.

The proposed model was compared with some other classification models: DenseNet121, EfficientNetB4, ResNet101, VGG16, and InceptionNetV3. It achieved the following results on the testing set: AUC – 0.8012, specificity – 0.7452, accuracy – 0.7429, precision – 0.2627, F1 – 0.3554. Taking into consideration the achieved results, there can be concluded, that application of a Vision Transformer with convolutional layers can give more accurate results than pure classification models, which are frequently used in this field of research. Additionally, the application of ASL can improve the accuracy of predictions, since this loss function is proposed for multilabel classification.

In future work, the model can be extended by experimenting with different optimizers, loss functions and modifications of the Vision Transformer. Additionally, some other techniques in preprocessing step can increase the accuracy of disease detection.

TABLE II
RESULTS FOR ALL METHODS.

Method	AUC	Sensitivity	Specificity	Accuracy	Precision	F1
DenseNet121	0.7667	0.7169	0.6965	0.7022	0.2324	0.3234
EfficientNetB4	0.7772	0.7079	0.7221	0.7223	0.2419	0.3324
ResNet101	0.7520	0.7011	0.6917	0.6975	0.2320	0.3232
VGG16	0.7797	**0.7427**	0.6930	0.6978	0.2433	0.3357
InceptionNetV3	0.7522	0.7073	0.6869	0.6864	0.2314	0.3171
Proposed	**0.8012**	0.7272	**0.7452**	**0.7429**	**0.2627**	**0.3554**

TABLE III
ROC-AUC FOR ALL METHODS.

Pathology	DenseNet121	EfficientNetB4	ResNet101	VGG16	InceptionNetV3	Proposed
Atelectasis	0.7325	0.7421	0.7450	0.7587	0.7379	**0.7635**
Cardiomegaly	0.8513	0.8594	0.8518	0.8684	0.8423	**0.8825**
Consolidation	0.6635	0.6663	0.6665	0.6890	0.6614	**0.6907**
Edema	0.8208	0.8187	0.8128	0.8253	0.7902	**0.8310**
Effusion	0.7886	0.7858	0.7944	0.8058	0.7815	**0.8114**
Emphysema	0.8313	0.8490	0.8088	0.8673	0.8147	**0.8988**
Fibrosis	0.7921	**0.8263**	0.7983	0.7888	0.7999	0.8213
Hernia	0.8549	0.8910	0.6222	0.7860	0.7138	**0.9065**
Infiltration	0.6791	0.6850	0.6762	0.7035	0.6727	**0.7054**
Mass	0.7797	0.7803	0.7865	0.8073	0.7773	**0.8098**
Nodule	0.7306	0.7442	0.7427	0.7541	0.7216	**0.7692**
Pleural Thickening	0.7011	0.7222	0.7181	0.7258	0.7138	**0.7666**
Pneumonia	0.6632	0.6644	0.6594	0.6786	0.6684	**0.6926**
Pneumothorax	0.8446	0.8461	0.8455	0.8576	0.8346	**0.8671**

REFERENCES

[1] S. Suganyadevi, V. Seethalakshmi, and K. Balasamy, "A review on deep learning in medical image analysis," *International Journal of Multimedia Information Retrieval*, vol. 11, no. 1, pp. 19–38, 2022.

[2] N. Subramanian, O. Elharrouss, S. Al-Maadeed, and M. Chowdhury, "A review of deep learning-based detection methods for covid-19," *Computers in Biology and Medicine*, p. 105233, 2022.

[3] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2097–2106, 2017.

[4] K. Nair, A. Deshpande, R. Guntuka, and A. Patil, "Analysing x-ray images to detect lung diseases using densenet-169 technique," *Available at SSRN 4111864*, 2022.

[5] E. Khan, M. Z. U. Rehman, F. Ahmed, F. A. Alfouzan, N. M. Alzahrani, and J. Ahmad, "Chest x-ray classification for the detection of covid-19 using deep learning techniques," *Sensors*, vol. 22, no. 3, p. 1211, 2022.

[6] K. Chen, X. Wang, and S. Zhang, "Thorax disease classification based on pyramidal convolution shuffle attention neural network," *IEEE Access*, vol. 10, pp. 85571–85581, 2022.

[7] V. Ravi, V. Acharya, and M. Alazab, "A multichannel efficientnet deep learning-based stacking ensemble approach for lung disease detection using chest x-ray images," *Cluster Computing*, pp. 1–23, 2022.

[8] F. J. M. Shamrat, S. Azam, A. Karim, K. Ahmed, F. M. Bui, and F. De Boer, "High-precision multiclass classification of lung disease through customized mobilenetv2 from chest x-ray images," *Computers in Biology and Medicine*, p. 106646, 2023.

[9] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[10] G. I. Okolo, S. Katsigiannis, and N. Ramzan, "Ievit: An enhanced vision transformer architecture for chest x-ray image classification," *Computer Methods and Programs in Biomedicine*, vol. 226, p. 107141, 2022.

[11] M. M. Al Rahhal, Y. Bazi, R. M. Jomaa, A. AlShibli, N. Alajlan, M. L. Mekhalfi, and F. Melgani, "Covid-19 detection in ct/x-ray imagery using vision transformers," *Journal of Personalized Medicine*, vol. 12, no. 2, p. 310, 2022.

[12] X. Jiang, Y. Zhu, G. Cai, B. Zheng, and D. Yang, "Mxt: A new variant of pyramid vision transformer for multi-label chest x-ray image classification," *Cognitive Computation*, vol. 14, no. 4, pp. 1362–1377, 2022.

[13] T. Wang, Z. Nie, R. Wang, Q. Xu, H. Huang, H. Xu, F. Xie, and X.-J. Liu, "Pneunet: deep learning for covid-19 pneumonia diagnosis on chest x-ray image analysis using vision transformer," *Medical & Biological Engineering & Computing*, pp. 1–14, 2023.

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

[15] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[19] A. Mezina and R. Burget, "Obfuscated malware detection using dilated convolutional network," in *2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 110–115, IEEE, 2022.

[20] T. Ridnik, E. Ben-Baruch, N. Zamir, A. Noy, I. Friedman, M. Protter, and L. Zelnik-Manor, "Asymmetric loss for multi-label classification," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 82–91, 2021.

Jsme skupina E.ON, těší nás

Jsme jedním z největších energetických koncernů u nás i ve světě. Centrálu máme v Německu, ale najdete nás už v 15 zemích Evropy, včetně České republiky. Zakládáme si na tom, že naše energie můžou šetřit peníze i přírodu a čím dál víc využíváme obnovitelné zdroje.

Kdo patří do naší rodiny

E.ON Energie, a.s.
- Obchoduje s elektřinou a plynem, zajišťuje marketing a komunikaci a stará se i o výrobu energií.
- Pro zákazníky připravuje řešení na míru v oblasti fotovoltaiky, tepelné techniky a elektromobility.

EG.D, a.s.
- Distributor energií, který vlastní a provozuje rozvodnou síť elektřiny zejména na jihu Čech a Moravy a rozvodnou síť plynu na jihu Čech.
- Zajišťuje připojení odběrných míst k síti a stará se o dopravu energií k zákazníkům.

E.ON Česká republika, s. r. o.
- Funguje jako podpora výše uvedeným společnostem. Zajišťuje jim služby, jako je účetnictví, právo nebo HR, a na starost má i zákaznickou péči.

Na přírodě nám záleží

E.ON Energie je jedničkou ve výkupu zelené elektřiny v České republice.

EG.D pomocí bezpečnostních prvků na sloupech vysokého napětí chrání ptáky před úrazem.

V budovách E.ONu i ve všech dobíječkách elektromobilů využíváme zelenou energii.

Nastartuj svou kariéru

#spolujsmeeon

Studentské programy
- Letní energetická akademie – 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

The Precise Measurement of Magnetic Azimuth

Veronika Junasová
Department of Microelectronics
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, Czech Republic
xjunas00@vutbr.cz

Vladimír Levek
Department of Microelectronics
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, Czech Republic
levek@vut.cz

Pavel Šteffan
Department of Microelectronics
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, Czech Republic
steffan@vut.cz

***Abstract*—This article deals with measuring magnetic azimuth using inertial sensors. It introduces the steps in this process and obstacles that need to be solved in order to achieve more precise results. The magnetic azimuth can be determined using one magnetometer; however, the precision in such measurement is limited. Therefore, this article also introduces measurement of relative azimuth acquired by the gyroscope. The essential part of precise measurement is the calibration of the mentioned inertial sensors, which is described in detail. The process of calibration is supported with equations, graphs, and tables. The measurement itself consists of an output from the magnetometer. The relative azimuth acquired from the gyroscope is then combined using a complementary digital filter with magnetic azimuth from more accurate magnetometer. The final results show the improvement obtained with the presented method. The article also contains information on a compact device designed for simple calibration and measurement with the mentioned sensors.**

Keywords*—*magnetic azimuth, relative azimuth, magnetometer, gyroscope, calibration

I. Introduction

The magnetic azimuth represents the tilt from north, in degrees. This information is used, for example, in aviation, telephones, space applications, or geographic applications [1]. Magnetic azimuth can be obtained relatively easily from one magnetometer. From the output of the magnetometer, it is possible to determine the magnetic azimuth with the use of trigonometric functions. However, this is not the only step in the process, as acquired values would not make sense this way.

The relevance of measuring magnetic azimuth depends on several factors. These are effects of surroundings, accuracy of calibration, uncertainty of measuring sensor. The surroundings of magnetometer consist of forces that are desirable for the particular application and of forces caused by other magnetic materials that generate magnetic field on its own. Therefore, the magnetometer measures the magnetic field surrounding it, but with unwanted fractions. These segments need to be eliminated during calibration, otherwise they can cause drift of measured results. Using the right calibration, the undesirable effects of surroundings can be avoided and they became non-essential.

The uncertainty of the magnetometer depends on the specific type of sensor, which therefore has to the be chosen in relevance to specific requirements of the given application. However, this means that the results depend on these uncertainties and cannot be improved anymore with calibration, unless some other means of measuring are incorporated into the process. For such purpose, it is possible to use gyroscope due to its complementary features. It is possible to obtain relative azimuth from gyroscope output with the use of integration. The precision of measurement depends on calibration, uncertainty of gyroscope, and precision of derivation.

The relative azimuth alone does not bear any significance in determining the tilt from north. However, with the use of a proper digital filter, it can improve the precision of magnetic azimuth gained from the magnetometer. For the purpose of this specific application, the linear complementary filter was chosen. It allows one to combine results in the same units from two different sensors with complementary properties. This paper covers in detail the issue of calibration and measurement of azimuth with magnetometer and gyroscope. It also presents device that was measured for such measurement.

II. Methods of Calibration and Measurement

A. Calibration of Magnetometers

The magnetometer is a device capable of measuring a permanent or temporal magnetic field surrounding it. The output of the magnetometer provides information about the direction and force of the magnetic field. From this the magnetic azimuth can be determined with the use of trigonometric functions. However, before the initiation of measurement, the calibration should be executed, otherwise the data would not be relevant for further processing. During calibration, the influence of unnecessary parts of magnetic field is removed. The calibration should take place at the exact same spot where the measurement will take place. The sensor environment contains hard and soft magnetic materials, from which calibration is called hard ironing and soft ironing [2]. The calibration here is performed in two-dimensional space.

If the magnetometer is in the place without interference, it observes the same absolute value of magnetic force in all directions with the centre in zero. Therefore, in an ideal situation without the effect of hard and soft magnetic materials, the data measured by magnetometer create a perfect circle with the centre at the origin of the coordinate system set by the coordinate x and y [3]. This is the desired shape of the measured values that is to be obtain after calibration.

The hard and soft magnetic materials have different effect on the shape of the circle. The hard magnetic materials create a permanent magnetic field around them, which causes the constant offset from the origin of the coordinate system. That means that the circle still has the same shape; however, its centre is shifted. In order to eliminate this effect, the constant value of offset has to be subtracted from measured values in each axis. This offset is specific for given axis and therefore must be determined separately [3].

The process of hard ironing, respectively, the removal of constant additive distortion caused by hard magnetic materials, consists of turning the sensor around one axis at least once. For each axis, the maximum and minimum must be found. The offset is given as average of the maximum and minimum values of the axis. The example of an equation for the axis x is in:

$$\mathrm{x_{Mag\_offset}} = \frac{\mathrm{x_{Mag\_max}} + \mathrm{x_{Mag\_min}}}{2}, \tag{1}$$

where x_{Mag_offset} is permanent offset on axis x, x_{Mag_max} is maximum value, and x_{Mag_min} is minimum value on axis x. This offset is then subtracted from the new measured values from the magnetometer. This way the centre of circle can be shifted to the origin of coordinate system, as desired.

The soft magnetic materials can influence the magnetic field around them, even though they do not create their own field. They flatten the circle and warp it. The circle then changes its shape to an ellipse. In application like this project the warp is undesirable and distortion is the main part that has to be removed. To eliminate this, the distortion factor needs to be determined. This factor is different for each axis, similarly to hard ironing process. It is better to initiate soft ironing after hard ironing, because the centre would be in the origin.

Soft ironing works with values that have undergone hard ironing. The same process must be done on every axis. From the maximum and minimum in the axis x their delta is determined, as in:

$$\mathrm{x_{Mag\_delta}} = \frac{\mathrm{x_{Mag\_HI\_max}} - \mathrm{x_{Mag\_HI\_min}}}{2}, \tag{2}$$

where x_{Mag_delta} is the desired delta, $x_{Mag\_HI\_max}$ is the maximum value on the axis x after hard ironing, and $x_{Mag\_HI\_min}$ is the minimum value on the axis x after hard ironing. The next step requires the knowledge of this delta for all used axes. Therefore y_{Mag_delta} in the following equation is the desired delta for the axis y. From deltas, the average avg_{Mag} has to be enumerated as:

$$\mathrm{avg_{mag}} = \frac{\mathrm{x_{Mag\_delta}} + \mathrm{y_{Mag\_delta}}}{2}. \tag{3}$$

The distortion factor is determined for the axis x and the axis y separately and is given as:

$$\mathrm{f_{Mag\_x}} = \frac{\mathrm{avg_{mag}}}{\mathrm{x_{Mag\_delta}}}, \tag{4}$$

where f_{Mag_x} is the distortion factor specific for axis x. The new values that have undergone hard ironing are multiplied by the distortion factor for the given axis. The whole calibration against both hard and soft magnetic materials is then:

$$\mathrm{x_{Mag\_cal}} = \left(\mathrm{x_{Mag\_new}} - \mathrm{x_{Mag\_offset}}\right) \cdot \mathrm{f_{Mag\_x}}, \tag{5}$$

where x_{Mag_cal} is the fully calibrated x-value and x_{Mag_new} is the newly measured x-value from magnetometer.

B. *Calibration of Gyroscopes*

The gyroscope is an inertial sensor that can measure the change in angular velocity. From this change in angular velocity, it is possible to acquire the change in angle using integration [2]. However, this brings additional problems caused by the drift of the gyroscope. Due to the integration, even small drift or noise of the device propagate itself significantly in time. For these reasons, it is impossible to use gyroscope for determination of azimuth on its own, and it is usually combined together with other inertial sensors. The calibration has to be completed before the actual measurement, the same as it is with magnetometers.

Gyroscope calibration is simple and does not require as much mathematical processing as is needed with magnetometers. During the calibration the gyroscope should be laid still on a flat surface. This way there is no change of forces that would affect the gyroscope. Therefore, the values in the gyroscope output should ideally be zero. However, these data show certain value of the change in angular velocity. These values present the offset from zero. The calibration consists of taking a sufficient number of samples and enumerating the average from them for each axis individually, as in:

$$\mathrm{x_{Gyr\_offset}} = \frac{\sum_{\mathrm{k=1}}^{\mathrm{i}} \mathrm{x_{Gyr\_k}}}{\mathrm{i}}, \tag{6}$$

where x_{Gyr_offset} is the offset in axis x, k is the index of measured sample, i is the number of samples and x_{Gyr_k} is the measured sample in axis x. The offset is then subtracted from newly measured data, and the calibrated values are given as:

$$x_{Gyr\_cal} = x_{\mathrm{Gyr\_new}} - x_{\mathrm{Gyr\_offset}}, \tag{7}$$

where x_{Gyr_cal} is the calibrated x-axis value and x_{Gyr_new} is the newly measured value on the axis x. The same process should be applied to all axes. The values after calibration will still not be perfect zero due to the noise of the device.

C. *Measurement of Magnetic Azimuth*

The magnetic azimuth obtained from the magnetometer provides information about the north. For the purpose of this paper, the magnetic azimuth is defined in the plane given by the x and y axis. The calibrated output value of the magnetometer is labelled as M_x for the axis x and M_y for the axis y. The magnetic azimuth acquired from magnetometer is labelled as ψ_{Mag}; this is the sign of the azimuth around the axis z commonly used in the literature [4]. The value of magnetic azimuth can be determined as [3]:

$$\psi_{\mathrm{Mag}} = \tan^{-1}\left(\frac{\mathrm{M_x}}{\mathrm{M_y}}\right) \ [°]. \tag{8}$$

The arctangent function of (8) gives output values from −90° to 90°. However, this is not sufficient range for representation of the whole circle around the axis z, respectively, from north. In order to fully express the tilt from the north in degrees, the range needed is from −180° to 180°. Moreover, the zero value of M_y is not defined in this function. The adjusted equation is the solution for the problems mentioned:

$$\psi_{\mathrm{Mag}} = \mathrm{atan2}\left(\frac{\mathrm{M_x}}{\mathrm{M_y}}\right). \tag{9}$$

The result from (9) is in radians and needs to be converted to degrees. The range now is from −180° to 180°, as desired. This range can be further optimised using simple addition or subtraction, for example, from 0° to 360°. Similarly, the north can be set in any desired angle and other directions would be recounted accordingly.

D. Measurement of Relative Azimuth

The azimuth obtained from the gyroscope is relative. That means that it gives information about angle from a certain starting point. Without the information about this staring point, the angle does not make sense alone for the detection of north or the tilt from north. The calibrated output value of the gyroscope is labelled as G_z for the axis z and the relative azimuth acquired from the gyroscope is ψ_{Gyr}. The relative azimuth can be calculated with a simple use of integration. This means that the newly measured value from the gyroscope can be multiplied by the time passed from the previous measurement and added to the angle from the previous measurement as:

$$\psi_{\mathrm{Gyr}} = \psi_{\mathrm{prev}} + \mathrm{G_z} \cdot \Delta\mathrm{t}, \tag{10}$$

where Δt is the time passed between two measurements. There are several methods of integration that can result in different precision of measurement. However, the difference in this application would be insignificant.

E. Digital Filter

The outputs from magnetometer and gyroscope can be combined with the use of an appropriate digital filter. The ideal filter for combining these two is the linear complementary filter. It incorporates the high-pass and low-pass filter [5]. There the complementary properties of magnetometer and gyroscope are advantageous. Magnetometers tend to have a better response to slowly changing signals. However, when the signal changes quickly, it can have a false reading. On the other hand, the gyroscopes have better response to quickly changing signals. When the change of signal is slow, the force applied to gyroscope is smaller and the measurement is more accurate. This means that they have complementary dynamic responses and are ideal for the usage in linear complementary filter [5]. Mathematically, this filter is given as:

$$\psi_{\mathrm{Filt}} = \alpha \cdot \psi_{\mathrm{Gyr}} + (1-\alpha) \cdot \psi_{\mathrm{Mag}}, \tag{11}$$

where ψ_{Filt} is the magnetic azimuth obtained from the application of the linear complementary filter and α is a weighted factor with value in the range <0; 1>.

III. Results

A. Designed Device

The device designed for the measurement of different movements in space, including magnetic azimuth, contains two magnetometers, an inertial measurement unit (IMU), and processing microcontroller. The magnetometers used are MMC5633NJL [6], further marked MMC, and BM1422AGMV [7], further marked BM. The IMU IIM-42652 [8] contains an accelerometer and gyroscope, although though in this paper only gyroscope is used. The device mounted on the base for calibration is depicted in Fig. 1.

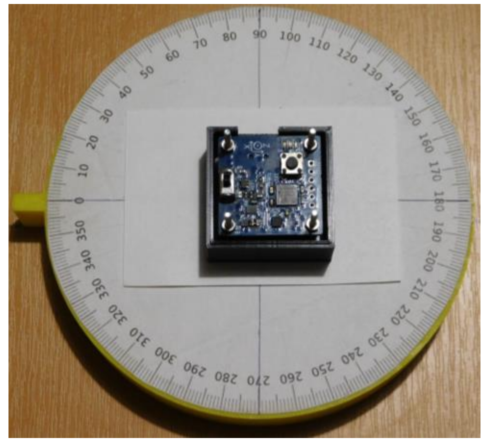


Fig. 1. Designed device mounted on the calibration base.

The device can operate with battery as power supply while saving measured data to its external memory. However, this operation is restricted due to the limited space in memory. In the second type of operation, the device is connected to the computer, which is also the power supplier at that moment. The device communicates through a terminal. The user can send prepared commands and the device executes the appropriate function. There are functions for calibrations of all present sensors and for measurement. Every calibration lasts a given time until it ends. The measurement functions continue until the appropriate command is received.

B. Calibration of Sensors

In the case of this particular application, the calibrated magnetometer was turned 360° around axis z. The data in this part are from the calibration of magnetometer MMC. The output data from this magnetometer before calibration can be seen in Fig. 2. They were obtained with the same motion as is done during calibration. It is obvious that the centre is not in the origin of coordinate system and that the shape is not circular.

The Fig. 3 shows the output data after hard ironing with blue colour and after soft ironing with orange colour. The hard ironing helped to move the centre of the circle to the origin of the coordinate system. The soft ironing then helped with the elliptical shape and converted it to the circle.

The offset and distortion of the circle before calibration is also obvious from the Table I. and Table II. In the first table are values of maximum, minimum, offset, and distortion before the calibration. In Table 2 are these values after calibration. It is shown that after calibration, the minimum and maximum have the same absolute value, their centre is almost in the origin, and their distortion was fixed.

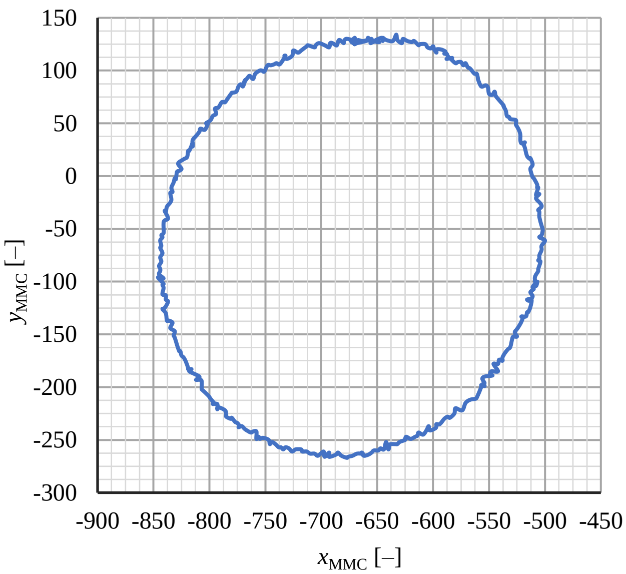


Fig. 2. The measured data from magnetometer MMC while turning around 360°around the axis z.

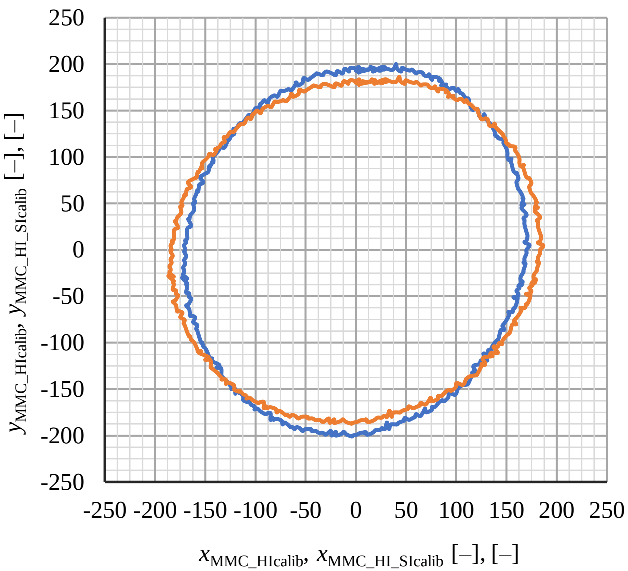


Fig. 3. The measured data from the magnetometer MMC after hard ironing (blue colour) and soft ironing (orange colour).

TABLE I. MAXIMUM, MINIMUM, OFFSET, AND DISTORTION BEFORE CALIBRATION

	Maximum	Minimum	Offset	Distortion
x_{MMC} [-]	-500	-846	-673	1,0748
y_{MMC} [-]	134	-267	-66	0,9314

TABLE II. MAXIMUM, MINIMUM, OFFSET, AND DISTORTION AFTER CALIBRATION

	Maximum	Minimum	Offset	Distortion
x_{MMC_cal} [-]	186,75	-186,75	0	1
y_{MMC_cal} [-]	186,28	-187,22	-0,47	1

The gyroscope was places during calibration intact on the flat surface and 500 samples were collected. Fig. 4 depicts the data before and after calibration. The blue colour is for the axis x before calibration, the orange colour is for the axis y before calibration, the red colour is for the axis x after calibration, and the yellow colour is for the axis y after calibration. This also supports Table III. with values of offsets before calibration for the axes x and y. The values after calibration are accumulated around zero.

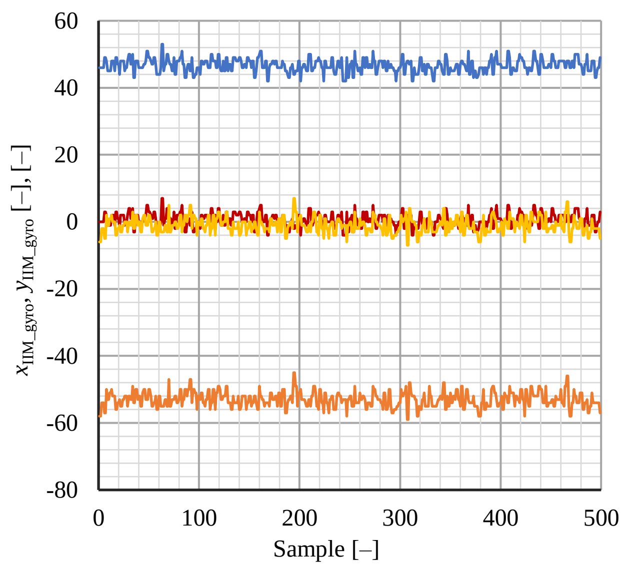


Fig. 4. Measured data from magnetometer gyroscope before calibration for the axis x (blue colour), axis y (orange colour), and after calibration for axis x (red colour), axis y (yellow colour).

TABLE III. THE OFFSET AND DISTORTION BEFORE CALIBRATION

	Offset
x_{IIM_gyro} [-]	46
y_{IIM_gyro} [-]	-52

C. *Magnetic Azimuth Measurement*

To identify the precision of measurement by magnetometer MMC alone, the following method was proposed. The device was set at defined angles from 0° to 360°. It was set with a 10° step and every time the magnetic azimuth was measured. The measured values were then compared to the set angle, which was considered ideal. The acquired data, marked as $azimuth_{mag}$, were plotted in relation to the ideal set angle, labelled as $azimuth_{set}$. The mentioned graph is shown in Fig.5 in which the blue colour is the measured azimuth and the orange colour means the ideal theoretical value. It is evident that the magnetic azimuth deviates from ideal angles, more visibly in the higher range of angles.

The improvement of this can be seen in Fig. 6. There is measured magnetic azimuth, marked as $azimuth_{filt}$ obtained after application of complementary filter. The measurement has the same conditions as the measurement of magnetometer alone. The device was set from 0° to 360° angles with 10° step, and magnetic azimuth is acquired. From Fig. 6 it is evident that after filtration the magnetic azimuth follows the ideal angles more precisely.

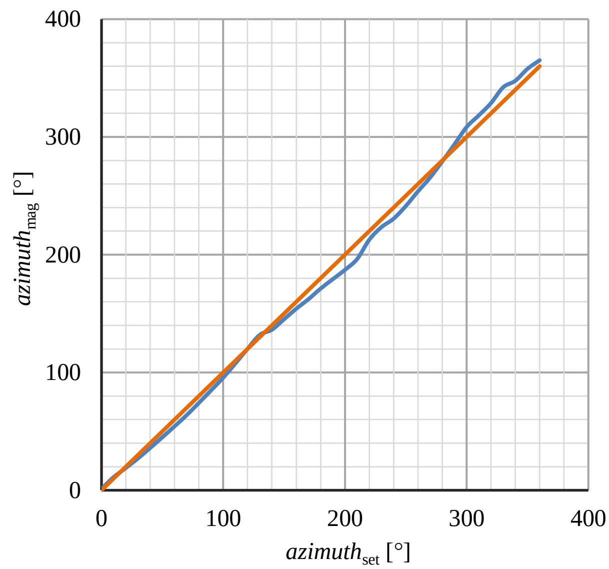


Fig. 5. The magnetic azimuth measured by magnetometer MMC in relation to the set angle (blue colour), the ideal angles (orange colour).

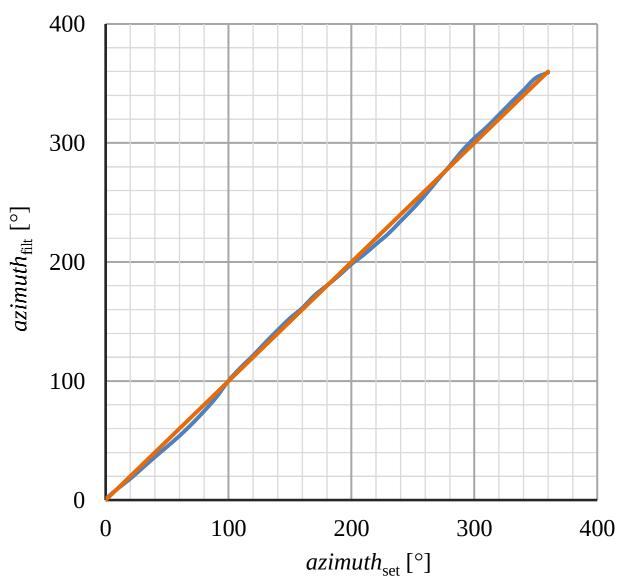


Fig. 6. The magnetic azimuth measured by magnetometer and gyroscope after application of digital filter in relation to the set angle (blue colour), the ideal angles (orange colour).

The values of magnetic azimuth after the application of the filter overlap more with theoretical values. Therefore, the magnetic azimuth with the use of filter is closer to the desired precision. This proves that the linear complementary filter is helpful for more precise measurement of magnetic azimuth. The accuracy could be further improved with the use of external crystal, which can potentially improve the accuracy of integration used with gyroscope. Furthermore, a different type of digital filter could bring new possibilities to this area.

IV. Conclusion

This article describes the process of measuring magnetic azimuth and proposes the use of digital filter in order to obtain more precise results. It includes detailed information concerning calibration of magnetometers and gyroscopes. There are all necessary equations for this purpose, and calibration of both sensors is supported with graphs and tables. The article also presents the device that was used for the purpose of all calibrations and measurements, and some of its properties were stated. The measurement of magnetic azimuth was done by magnetometer alone and using magnetometer and gyroscope combined by filter. The measurement proved that the combination of inertial sensors helps with the precision of the magnetic azimuth. Further improvement in this field could be obtained with improvement of integration used in determination of relative azimuth, the use of several magnetometers in one measurement or the usage of different types of filtering.

Acknowledgment

The article was supported by project no. FEKT-S-23-8162, Modern micro- and nanoelectronics for future.

References

[1] J. FRADEN, Handbook of Modern Sensors: Physics, Designs, and Applications, 5th ed, San Diego: Springer International Publishing, 2016. ISBN 978-3-319-19302-1.

[2] Á. López Revuelta, “Orientation estimation and movement recognition using low cost sensors,” Blekinge Institute of Technology, Karlskrona, Sweden, pp. 79, 2017.

[3] Ch. KONVALIN, “Compensating for Tilt, Hard-Iron, and Soft-Iron Effects,” Questex, 2009.

[4] T. YOUNG, Performance of the Jet Transport Airplane: Analysis Methods, Flight Operations, and Regulations, 1st ed, Wiley, 2018. ISBN 978-1-119-68278-3.

[5] P. Narkhede, S. Poddar, R. Walambe, G. Ghinea and K. Kotecha, “Cascaded Complementary Filter Architecture for Sensor Fusion in Attitude Estimation,” Sensors, no. 6: 1937, 2021.

[6] MEMSIC, “±30 Gauss, Monolithic, High Performance, Low Cost 3-axis Magnetic Sensor: MMC5633NJL,” 2020.

[7] ROHM Semiconductor, “3-Axis Digital Magnetometer IC: BM1422AGMV,” 2016.

[8] TDK InvenSense, “IIM-42652 Datasheet: High-performance 6-Axis SmartIndustrial™ MotionTracking MEMS Device for Industrial Applications,” 2021.

Methods for the fast detection of water and degradation changes in lithium based electrolytes

Antonín Šimek
Department of Electrical and Electronic Technology
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
Antonin.Simek1@vut.cz

Ondřej Čech
Department of Electrical and Electronic Technology
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech republic
cechondrej@vut.cz

***Abstract*—This paper deals with lithium-ion batteries, specifically the electrolytes used. The degradation mechanisms of lithium-ion batteries are addressed, with a focus on electrolyte degradation. A considerable part is then devoted to the degradation of lithium hexafluorophosphate, especially after exposure to water or moisture. The decomposition of the lithium salt is linked to the evolution of hydrofluoric acid, whose influence on the electrolyte parameters and methods of fast detection is the subject of the experimental section.**

Keywords—lithium-ion batteries, electrolytes, degradation mechanisms, safety, water contamination, fast detection

I. Introduction

Over the past three decades, lithium-ion batteries have made a significant contribution to the energy storage systems. Lithium-ion batteries have become the preferred power source in consumer electronics and are continuously entering new market segments. At the same time this technology is rapidly growing in the field of electric vehicles and stationary energy storages. This is mainly due to their superior characteristics in terms of relatively high energy and power density (both volumetric and gravimetric), high energy efficiency, possibility of fast charging, adjustable battery size and sufficient safety features compared to other types of batteries. [1]

In general, lithium-ion batteries consist of a positive and negative electrode, an electrolyte, a separator and a casing. This system works on the intercalation principle, with the negative electrode most often made of graphite, and the positive electrode made of transition metal oxides such as Lithium Cobalt Oxide ($LiCoO_2$, or LCO), Lithium Nickel Manganese Cobalt Oxides ($LiNi_xMn_yCo_zO_2$ or NMC), Lithium Iron Phosphate ($LiFePO_4$, or LFP), and others [2]. These electrodes are separated by a porous polymer separator that prevents electrical contact between the two electrodes, but at the same time allows lithium ions to pass between the positive and negative electrode. The conductive pathway for the lithium ions is then formed by the electrolyte, which is generally a lithium salt dissolved in a suitable solvent or a mixture of solvents [2]. The composition of electrolytes for lithium-ion batteries is the main focus of this paper.

The limiting factors of a battery are its internal components. The electrodes, in particular the positive electrode, is the limiting factor in terms of overall battery capacity, energy density and cyclability. In contrast, the electrolyte determines the maximum current density, time stability and safety. This is due to the fact that the electrolyte is in close interaction with all other components such as the positive and negative electrodes as well as the separator. [3]

The interface between each electrode and the electrolyte determines the requirements for chemical stability, thereby greatly limiting the range of usable materials. The chemical compatibility of the electrodes with the electrolyte is ensured by the formation of passivation layers, referred to as Solid Electrolyte Interface (SEI for short). The formation and physical properties of these protective layers depend on the nature of the electrodes. Thus, more and more attention is being paid to the study of the electrolyte composition to ensure the stability and durability of the entire battery system. [3]

II. Lithium-ion Battery Degradation

Increased demands for lithium-ion batteries not only in automotive applications, such as early mentioned increased energy density, fast charging capabilities and improved safety, are accompanied by phenomena that negatively affect battery performance [1]. This so called degradation mechanisms (or aging) of lithium-ion batteries are shown on Fig 1 bellow. This figure shows the basic structural components of a lithium-ion battery and their associated degradation mechanisms. The primary degradation mechanisms are marked in green, the secondary mechanisms are marked in red [4].

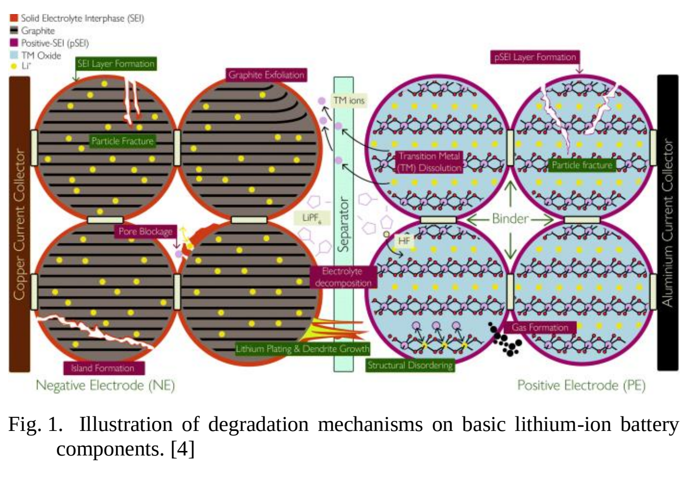


Fig. 1. Illustration of degradation mechanisms on basic lithium-ion battery components. [4]

As can be seen in Figure 1, the primary degradation mechanisms include particle cracking at the positive and negative electrodes, SEI formation and growth (at both positive and negative electrodes), structural changes of the positive electrode associated with dissolution of transition metal oxides, lithium plating of the negative electrode and associated dendritic growth. Secondary degradation mechanisms include binder decomposition associated with delamination of the active material from the current collector, formation of islands of active material due to particle cracking, pore blocking, evolution of gases and hydrofluoric acid, and electrolyte decomposition [5]. The decomposition of the electrolyte and its susceptibility to the presence of moisture in the presence of hydrofluoric acid will be the subject of one of the next chapter.

Battery degradation can generally be described from two perspectives. The first is to describe the degradation mechanisms in terms of the physical and chemical changes that take place inside the battery cell. These mechanisms describe degradation in the most detail, but at the same time are very difficult to observe directly. In contrast, the observation of power loss and usable capacity of a battery cell is relatively simple, but at the same time the least accurate for describing degradation mechanisms. However, understanding degradation mechanisms is crucial in designing cost-effective batteries with respect to long life and environmental sustainability. Therefore, there is a growing interest in the possibility of direct observation of degradation mechanisms during battery operation (in-situ measurement). [4]

III. Electrolytes for Lithium-ion Batteries

Electrolytes for lithium-based batteries in general can be divided in to five different groups:

- non-aqueous electrolytes made out of a lithium salt dissolved in an suitable organic solvent or mixture of solvents;
- aqueous solutions consisting of a lithium salt solubilized in water;
- ionic liquids from organic lithium salts;
- polymer electrolytes (both gel and solid polymers) and
- hybrid electrolytes. [3]

In this paper the attention will be mainly paid to non-aqueous electrolytes since most of the electrolyte used in commercial lithium-ion batteries are non-aqueous solutions [3]. Focus will be paid to basic requirements for these electrolytes, as well as the most commonly used materials for non-aqueous electrolytes. A considerable part will also be devoted to the already mentioned degradation of electrolytes, especially in connection with the evolution of hydrofluoric acid after exposure to the moisture.

A. *Electrolyte Requirements*

The electrolyte is therefore an internal part of the battery which must be sufficiently stable with respect to the cathode and anode material and should not undergo chemical changes during battery operation. The ideal electrolyte should therefore meet the following criteria [3]:

- It should be a good ionic conductor as well as an electrical insulator. A good ionic conductivity is necessary because of the easy transport of Li^+ ions between the anode and cathode during charging/discharging, and the insulation resistance is necessary to prevent self-discharge and short circuit.
- It should have a wide potential window meeting the range of working potentials of the anode and cathode material to avoid electrolyte decomposition at higher potentials.
- It should be inert to the other constructional parts of the battery cell so that, the separator does not dissolve or the battery cover does not corrode.
- It should be thermally stable over a wide range of temperatures so that it does not approach the boiling point or freezing point during operation.
- It must have low environmental toxicity and generally meet the conditions limiting environmental contamination.
- It should be made from sustainable materials. In other words, the electrolyte should not be made from low abundant elements, and the production process should have as least impact to the environment as possible. At the same time, the costs of preparing the raw materials and the total cost of electrolyte should be as low as possible. [3]

In addition to the requirements mentioned above, the electrolyte plays a very important role in initial cycling. During the first cycles, SEI is formed on the negative electrode and CEI (Cathode Electrolyte Interface) on the positive electrode [5]. This electrode-electrolyte interphases consists of salt/solvent-based inorganic and organic parts is electrically insulated while the ionic conductivity remains [1].

B. *Electrolyte Composition*

Lithium-ion batteries electrolytes are most often composed of a solution of Lithium Hexafluorphosphate (or $LiPF_6$) in mixture of carbon based solvents (or carbonates) [5].

These carbonates most commonly include Ethylene Carbonate (EC), Dimethyl Carbonate (DMC) and Diethyl Carbonate (DEC) [7]. Other carbonates and often used solvents are listed in Table 1. Mentioned solvents are used because of their acceptable stability for use in 4 V cathode systems and their stability to the lithiated graphite used for the negative electrode. Their other advantageous properties include high polarity, relatively good operating temperature range, sufficiently low toxicity and acceptable safety. [6]

TABLE I. BASIC PROPERTIES OF CARBONATE SOLVENTS [6]

Solvent	**Solvent Properties**					
	Formula	ε_r [-]	η [mPa.s]	ρ [g.cm-3]	***BP*** [°C]	***FP*** [°C]
Ethylene Carbonate (EC)	$C_3H_4O_3$	90	1.9	1.32	238	143
Propylene Carbonate (PC)	$C_4H_6O_3$	65	2.5	1.20	242	138
Dimethyl Carbonte (DMC)	$C_3H_6O_3$	3.1	0.59	1.06	90	17
Ethyl-Methyl Carbonate (EMC)	$C_4H_8O_3$	3.0	0.65	1.01	108	23
Diethyl Carbonate (DEC)	$C_5H_{10}O_3$	2.8	0.75	0.97	127	25

Table 1 above lists the most commonly used carbonates and their basic properties. Among the properties listed are the chemical formula, relative permittivity (ε_r), viscosity (η), density (ρ), boiling point (BP) and flash point (FP). All properties listed are for 25 °C except for ethylene carbonate, for which the properties are listed for 40 °C [3].

According to [6], the use of $LiPF_6$ salt is to some extent a compromise, as other commercially available lithium salts have too many disadvantages. For example, $LiAsF_6$ is poisonous; $LiClO_4$ is explosive; $LiBF_6$ exhibits significant problems at the negative electrode as it tends to react between the BF_4^- anion and the negative electrode surface; $LiSO_3CF_3$ exhibits too low ionic conductivity; and $LiN(SO_2CF_3)_2$ exhibits problems at the positive electrode as there is insufficient passivation, resulting in corrosion of the aluminium current collector. [6]

C. *Advanced Electrolyte Composition to Increase Battery Pefrormance and Safety*

In recent years, there has been a great effort to increase the overall power that a battery cell can deliver. One way to achieve this increase in performance is through the use of new solvents, lithium salts and other additives that can affect performance of the system [6]. Another much-discussed problem with lithium-ion batteries is their safety, especially fire safety. In some respects, these concerns are preventing further use of this type of technology [9]. However, according to [9], most of the fire risks are due to the unstable and highly flammable electrolytes composed of $LiPF_6$ and carbonate solvents.

1) *Solvents Perspective*

According to [6], new solvents include organic sulphur compounds such as Propylene Sulphite ($C_3H_6O_3S$), fluorine-substituted compounds and organo-phosphorous compounds. It is further argued in [8] that by using non-flammable phosphate compounds such as Dimethyl Methyl Phosphonate (DMMP), Triethyl Phosphate (TEP) and Trimethyl Phosphate (TMP) instead of conventional carbonate solvents, some degree of safety was achieved, but with an impact on performance. In fact, the poor reducing stability of these solvents leads to the decomposition of the solvents on the graphite electrode, resulting in a reduction in the system performance.

2) *Lithium Salts Perspective*

It has also been found that safety problems arise due to a thermally unstable salt, which is most commonly $LiPF_6$. Under unsuitable conditions, the heat and by-products generated by the decomposition of $LiPF_6$ accelerate a number of exothermic reactions. These reactions include the decomposition of the SEI layer, which leads to further reactions between the electrode and the electrolyte. During these reactions, oxygen evolution and heat release occur, which can lead to battery fires. Therefore, replacing the $LiPF_6$ salt with a more stable salt appears to be a solution to this problem. This was achieved by using Lithium bis(fluorosulfonyl)imide (or LiFSI for short) and Lithium bis(trifluoromethanesulfonyl)imide (or LiTFSI for short). [7]

The problem with these salts, however, is the lack of formation of passivation layers on the aluminium current collector, which leads to its corrosion at potentials higher than 3.8 V. However, it was found that corrosion of the current collector occurs only with weak LiTFSI solutions. In fact, when a higher concentration of this salt was used, the collector was passivated by a layer of Aluminum Fluoride (AlF_3), which suppresses surface corrosion. Moreover, by increasing the concentration of LiTFSI, the proportion of flammable solvents is reduced, which increases the safety of the batteries. [7]

IV. Electrolyte Degradation and Decomposition

The aforementioned $LiPF_6$ dissolved in carbonate solvents shows a good compromise between operating temperature, ionic conductivity, SEI formation and potential window. However, the lack of redox stability leads to degradation mechanisms that manifest themselves in electrolyte decomposition. Electrolyte aging can be divided into:

- contribution to interphase formation;
- solvent decomposition and
- salt decomposition. [1]

A. *$LiPF_6$ Decomposition*

Although the most commonly used lithium salt, $LiPF_6$, is relatively stable in a dry inert atmosphere up to 107 °C, this salt suffers from degradation when exposed to traces of water, moisture, or even alcohols [8].

In [8] it is shown that the degradation starts with the decomposition of $LiPF_6$ according to equation (1). According to this equation, $LiPF_6$ decomposes to form Lithium Fluoride (LiF) and Phosphorous Pentafluoride (PF_5).

$$LiPF_6 \rightarrow LiF + PF_6 \tag{1}$$

PF_5 can further react with water (H_2O) or moisture to release Hydrofluoric Acid (HF) and Phosphoryl Fluoride (POF_3). This process can be described by equation (2):

$$PF_5 + H_2O \rightarrow POF_3 + 2HF \tag{2}$$

The reaction described by equation (1) is almost negligible at room temperature. However, in an electrolyte environment, the interaction between PF_5 and solvent molecules (EC and DEC) can shift the equilibrium of the reaction to the right, resulting in increased HF formation. Decomposition is not only catalysed by elevated temperature or the presence of moisture, but is also activated by the release of protons(H^+) during LiF formation either in the SEI at the negative electrode or in the CEI at the positive electrode. [8]

$$Li^+ + HF \leftrightarrow LiF(s) + H^+ \tag{3}$$

The precipitation of LiF(s) described by equation (3) and therefore the lithium ion removal from the solvated system has no direct effect on the formation of HF. Released protons may further react with $LiPF_6$ molecules to form additional HF which is described by the equation (4) and (5). [8]

$$H^+ + PF_6^- \leftrightarrow H{-}F{-}PF_5 \tag{4}$$

$$H{-}F{-}PF_5 \rightarrow HF + PF_5 \tag{5}$$

The HF produced by the reactions described above dramatically affects battery life and performance [8]. Thus, the moisture content during the battery manufacturing process is one of the most important factors affecting the quality of both the electrolyte and the entire battery system. Indeed, even trace amounts of water can lead to significant deviations from normal behaviour. Figure 2 shows some of the undesirable chemical and electrochemical reactions, the most important being the degradation of the cathode, anode and electrolyte itself. [9]

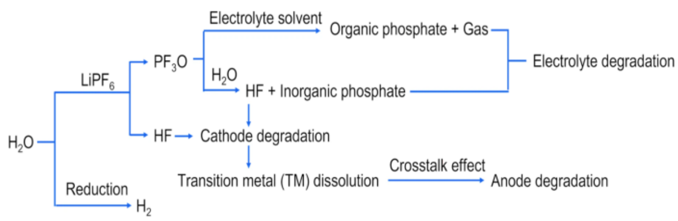


Fig. 2. Diagram showing the effect of water on the electrodes and electrolyte degradation [10]

V. Experimental

In the experimental part the potential window and stability of the electrolyte 1M $LiPF_6$ in EC:DMC 1:2 (i. e. 1 mol.l^{-1} of $LiPF_6$ dissolved in mixture of solvents composed from EC and DMC in 1:2 ratio) was tested. Furthermore, detection of the presence of water and degraded electrolyte was performed using fast electrochemical screening methods. For the potential window measurements, electrochemical test cells (ECC) from El-Cell were used. Their assembly and general handling of the electrolyte was carried out in a Jacomex glove box under the argon inert atmosphere (both H_2O and O_2 concentration bellow 1 ppm). Electrochemical measurements were then performed on a BioLogic VMP3 multi-channel potentiostat.

Potential window and stability measurements were performed using four different test cell configurations. Each configuration is used to test stability against specific battery components:

- Steel-Steel configuration (empty test cell with Whatman glass separator) to measure the stability of the electrolyte against the applied potential and when in contact with the separator and metal parts of the battery.
- Li-Steel configuration (lithium metal on one side, empty measuring cell on the other, Whatman glass separator in between) for measuring the stability of the electrolyte in the presence of lithium metal and Li^+ ions.
- Li-Cu configuration (lithium metal on one side, copper disk on the other, Whatman glass separator in between) for measuring the stability of the electrolyte against the copper current collector.
- Li-Al configuration (lithium metal on one side, aluminium disc on the other, Whatman glass separator in between) for measuring the stability of the electrolyte against the aluminium current collector.

As part of the electrochemical measurements, the open circuit voltage (OCV for short) was first measured for one hour. Subsequently, a set of voltametric measurements were performed in which the current response to the applied potential was measured. As part of these measurements, linear sweep voltammetry (LSV for short) and twenty cycles of cyclic voltammetry (CV for short). The scan rate of these methods was chosen to be 5 mV.s^{-1}. Potential range was chosen from 0.05 to 5.5 V for the Steel-Steel and Li-Steel configurations, 0.05 to 2.5 V for the Li-Cu configuration and 2.5 to 5.5 V for the Li-Al configuration. These potential ranges were chosen based on consultation with the supervisor, taking into account the stability of the materials used.

VI. Results and Discussion

This section will discuss the measured results on the above mentioned test cell configurations. All measurements were carried out with a three-electrode configuration. Since there were many test configurations and a lot of obtained data, only a comparison between pure and H_2O contaminated electrolyte will be presented here.

A. *Difference in obtained CV curves between pure and contamined electrolyte*

Figure 3 shows the CV curves obtained from measurements on the Li-Al configuration. The orange colour corresponds to the pure electrolyte, the blue colour corresponds to the electrolyte contaminated with 250 ppm H_2O. The red arrows then indicate what happened to the measured curves with increasing number of cycles. This colour coding is used for all the following results.

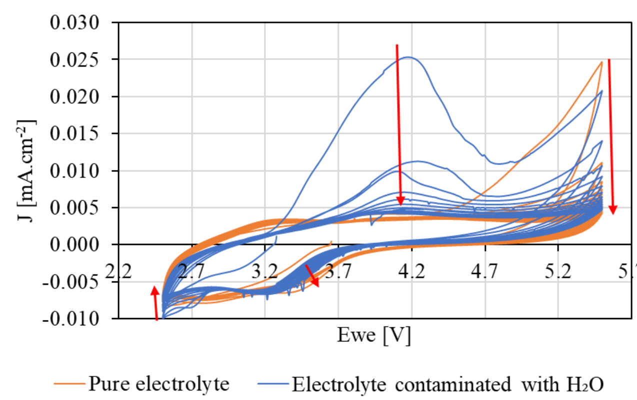


Fig. 3. CV curves of pure and contaminated electrolyte with 250 ppm H_2O on Li-Al configuration

From Figure 3 can be shown, that intense peak at a potential of approximately 4.2 V was present in the contaminated electrolyte. Furthermore, small micro-short circuits can be observed around potential of 3.2 V, which are most likely due to damage to the separator caused by etching of fibre glass separator.

Figure 4 shows the CV curves measured on the Li-Cu conjugation. Intense peaks at a potential of approximately 1.9 V can be observed in these curves which probably correspond to the electrolysis of H_2O. At lower potentials it is then possible to observe small spikes, which are probably due to damage to the separator caused by HF. Similarly to Figure 4, in Figure 5, where the curves for the Li-Steel configuration are plotted, it is possible to observe the appearance of peaks at a potential of approximately 1.5 to 2.0 V, which should also correspond to the electrolysis of H_2O.

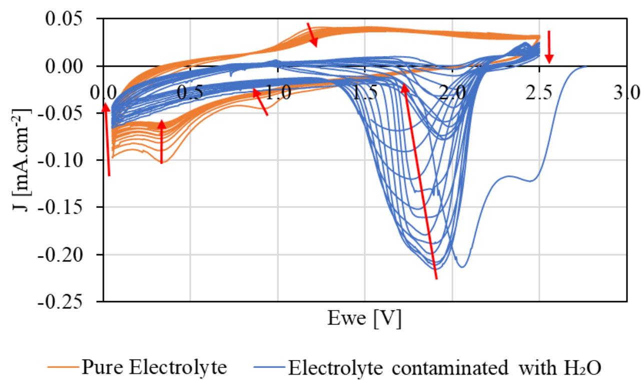


Fig. 4. CV curves of pure and contaminated electrolyte with 250 ppm H_2O on Li-Cu configuration

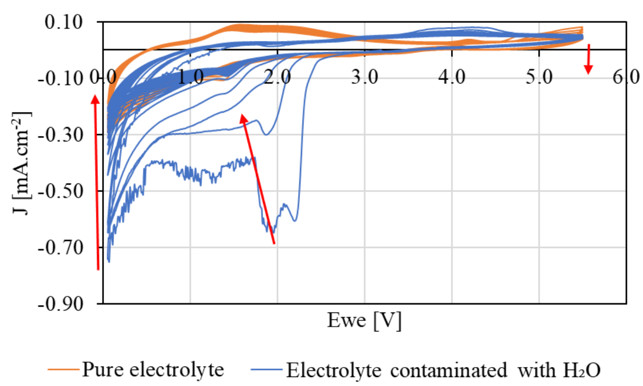


Fig. 5. CV curves of pure and contamined electrolyte withc 250 ppm H_2O on Li-Steel configuration

Figure 6 shows the results obtained from OCV measurements on a Li-Cu configuration using a water-contaminated electrolyte.

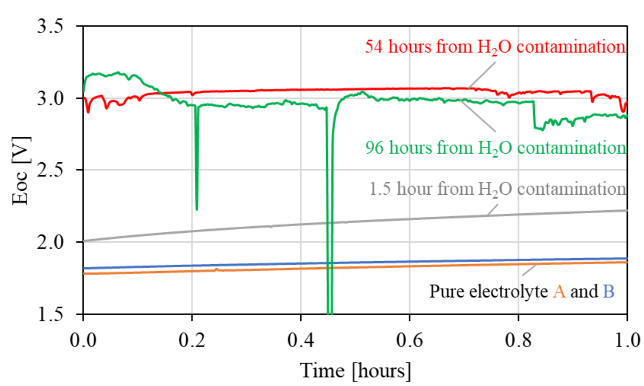


Fig. 6. OCV measurements on Li-Cu configuration with different time intervals from water contamination.

Figure 6 is based on two reference measurements on pure electrolyte, which are plotted in blue and orange. For both measurements, the OCV was around 1.85 V. Approximately 1.5 hours from water contamination, an OCV ranging from 2.0 to 2.25 V was measured, which is plotted in grey. After 54 hours, an OCV of around 3.0 V was measured, on which there are also small spikes caused by the breakdown of the separator. This measurement is plotted in red. After 96 hours from water contamination, the measured OCV was around 3.0 V, but with the difference that there are significant voltage drops caused by micro-short circuits through the HF-etched separator. These results show that the degradation of the electrolyte by the presence of water causes an increase in the open circuit voltage and also increases the HF concentration.

VII. Conclusion

The electrolytes for commercial lithium-ion batteries most often consist of $LiPF_6$ dissolved in carbonate solvents (such as EC or DMC). $LiPF_6$ has been found to be very sensitive to the presence of water and moisture, as it decomposes to form hydrofluoric acid when contaminated. This is why any contact with water and moisture during the battery manufacturing process must be avoided to ensure long-term reliability and service life. Usually the presence of water is detected by Karl-Fisher titration, which is a relatively complex and time-consuming method. During our research, rapid electrochemical screening methods were used, which can also detect the presence of water and degraded electrolyte.

Acknowledgment

This work was supported by the specific graduate research of the Brno University of Technology No. FEKT-S-23-8286.

References

[1] J. Henschel, F. Horsthemke, Y. P. Stenzel, M. Evertz, S. Girod, C. Lürenbaum, K. Kösters, S. Wiemers-Meyer, M. Winter, and S. Nowak, "Lithium ion battery electrolyte degradation of field-tested electric vehicle battery cells – A comprehensive analytical study", *Journal of Power Sources*, vol. 447, 2020.

[2] B. L. Ellis, K. T. Lee, and L. F. Nazar, "Positive Electrode Materials for Li-Ion and Li-Batteries", *Chemistry of Materials*, vol. 22, no. 3, pp. 691-714, Feb. 2010..

[3] Q. Li, J. Chen, L. Fan, X. Kong, and Y. Lu, "Progress in electrolytes for rechargeable Li-based batteries and beyond", *Green Energy & Environment*, vol. 1, no. 1, pp. 18-42, 2016.

[4] J. S. Edge, S. O'Kane, R. Prosser, N. D. Kirkaldy, A. N. Patel, A. Hales, A. Ghosh, W. Ai, J. Chen, J. Yang, S. Li, M. -C. Pang, L. Bravo Diaz, A. Tomaszewska, M. W. Marzook, K. N. Radhakrishnan, H. Wang, Y. Patel, B. Wu, and G. J. Offer, "Lithium ion battery degradation: what you need to know", *Physical Chemistry Chemical Physics*, vol. 23, no. 14, pp. 8200-8221, Apr. 2021.

[5] V. L. Martins, "Advances on liquid electrolytes for Li-ion and Li metal batteries", *Current Opinion in Electrochemistry*, vol. 38, 2023.

[6] D. Aurbach, Y. Talyosef, B. Markovsky, E. Markevich, E. Zinigrad, L. Asraf, J. S. Gnanaraj, and H. -J. Kim, "Design of electrolyte solutions for Li and Li-ion batteries: a review", *Electrochimica Acta*, vol. 50, no. 2-3, pp. 247-254, 2004.

[7] H. Liang, X. Zuo, L. Zhang, W. Huang, Q. Chen, T. Zhu, J. Liu, and J. Nan, "Nonflammable LiTFSI-Ethylene Carbonate/1,2-Dimethoxyethane Electrolyte for High-Safety Li-ion Batteries", *Journal of The Electrochemical Society*, vol. 167, no. 9, Apr. 2020.

[8] S. F. Lux, I. T. Lucas, E. Pollak, S. Passerini, M. Winter, and R. Kostecki, "The mechanism of HF formation in LiPF6 based organic carbonate electrolytes", *Electrochemistry Communications*, vol. 14, no. 1, pp. 47-50, 2012.

[9] X. Cao, "Important factors for the reliable and reproducible preparation of non-aqueous electrolyte solutions for lithium batteries", *Communications Materials*, vol. 4, no. 1, 2023

Recovery Of Anode Material From Lithium Ion Batteries

Jiří Báňa
Department of Electrical and Electronic Technology
Brno University of Technology
Brno, Czech Republic
0000-0002-8119-979X

Tomáš Kazda
Department of Electrical and Electronic Technology
Brno University of Technology
Brno, Czech Republic
0000-0003-1973-02Z

As the production of lithium-ion batteries (LIB) is growing at a meteoric pace with the advent of electromobility. There will be a need for efficient use of recycling processes, which now focus mainly on selected metals. Direct recycling is a method with great potential due to its simplicity, ecological friendliness, and high efficiency. However, this method needs to be improved and put into practical use. In this paper, it is focused the simulation of LIB aging, composition analysis and then recovery of the anode material of Motoma LFP 18650 and Panasonic NCR18650B batteries.

Keywords— Li-ion battery, direct recycling, graphite

I. INTRODUCTION

LIBs consist of a cathode material attached to an aluminum foil, an anode material on a copper foil, a separator that separates the cathode from the anode, an electrolyte that allows the migration of lithium ions, and everything is encapsulated in a metal package, most often made of aluminum. Graphite is most often used as the anode material due to its price and suitable properties. A number of materials are used as the cathode material, most often transition metal oxides and lithium. Cathode material can be chosen based on the required battery properties. [1] The total capacity on the market is expected to be more than 1.5 TWh in 2030, almost quadrupling compared to 2022. To the year 2040, total capacity of LIB on the market will attack 6 TWh. [2]

During the lifespan of LIBs, various mechanisms of degradation occur which also affect direct recycling. A common mechanism is the formation of an SEI layer. Another mechanism is particle cracking due to intercalation processes or lithium plating. [3]

Graphite dominates the commercial LIB market due to its excellent electrochemical properties and availability. The destruction of graphite occurs during cycling. Because of repeated deintercalation of Li^+ ions lead to a change of volume of the graphite anode. Cumulative stresses gradually cause graphite particles to crack and split. Typically, direct recycling strategies for graphite mainly focus on removing impurities and restoring its crystal structure.

The first step in repairing used graphite is to remove impurities, including SEI residue, polymer binder, conductive agents, contaminants, and metal deposits. Several processes are used to remove impurities, such as heat treatment, chemical treatment, acid treatment, and water treatment. Chemical treatment using ethylene glycol can not only remove the binder and conductive additives, but also dissolve metal impurities. Spray drying is then carried out to reconnect the graphite particles. Acid treatment is widely used to dissolve the metal. The performance of recycled graphite after acid treatment is close to commercial graphite. [4]

II. RECOVERY OF SPENT BATTERIES

Batteries used for this work are commercial Panasonic NCR18650B (3250 mAh) with NCA cathode and graphite anode and Motoma LFP 18650 (1500 mAh) with LFP cathode and graphite anode. Discharge characteristic was performed using different C-rates between 0,1C-1 C followed by long-term cycling to simulate aging of batteries. Cycling was performed by 1C-rate in the range declared by the manufacturer as safe for the cell for a total of 500 cycles. After aging, discharge characteristics were performed again. As shown in Figure 1 and Figure 2, capacity fade of Motoma LFP is minimal. However, capacity drop during 1 C cycling for Motoma LFP after 500 cycles is 5% as shown in Figure 3. The instruments used for electrochemical testing were a Biologic VMP3 with Booster and a ZKETECH EBC-X cycler.

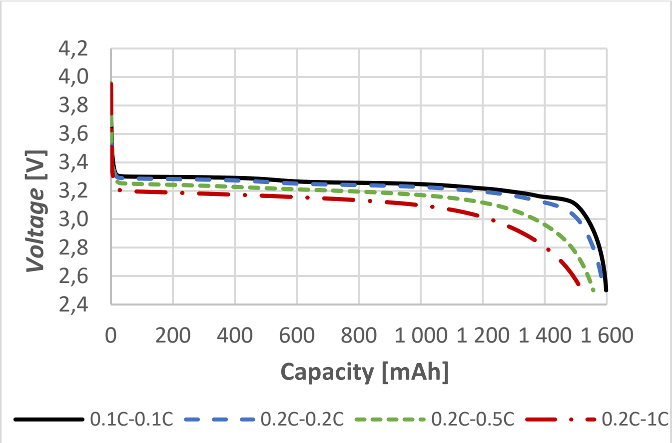


Figure 1 Discharging characteristics of Motoma LFP

Discharge curve of Motoma LFP before aging (Fig.1) is typical for LFP batteries. Initial drop of voltage followed by stable discharge plateau around 3.3 V. It was found that the capacity when discharging with a current of 0.1 C is 1590 mAh and at 0.2 C is 1560 mAh. The difference is only minor, and the voltage drop when discharging is also minor. At a load of 0.5 C the drop in the discharge plate, especially in the last third of the capacity, is more significant and the capacity reached was 1556 mAh. An even more significant drop in voltage occurred at a load of 1 C when the potential of the discharge plateau dropped below 3.2 V and the capacity dropped to 1519 mAh.

Figure 2 shows the discharge curves at different loads of the Motoma LFP battery after cycling. From the discharge characteristics it can be seen that there is a drop in capacity at all loads but at higher C there is not such a significant drop in voltage. So the battery behaves more stable at higher loads after cycling. The capacity at 0.1 C dropped by 9.4% percent to 1450 mAh, at 0.2 C the capacity was 1440 mAh. At 0.5 C the drop on the discharge plate after cycling, especially in the last third of the capacity, is less pronounced and the capacity reached was 1430 mAh. The drop in capacity at 1 C load was not as noticeable compared to the uncycled cell, the capacity dropped to 1410 mAh.

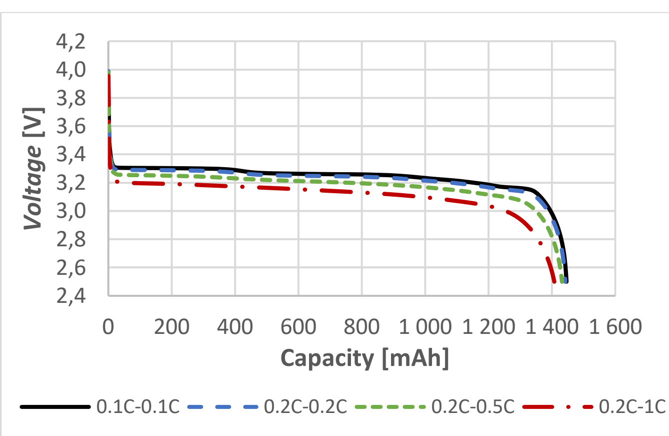


Figure 2 Discharging characteristics of aged Motoma LFP

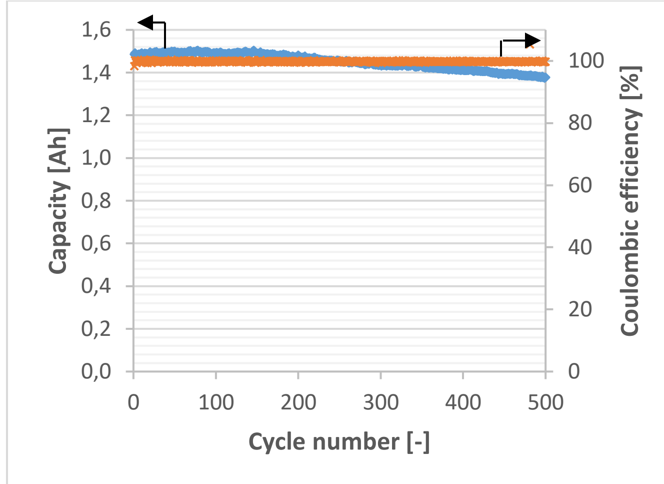


Figure 3 Capacity fade of Motoma LFP during cycling

Initial capacity during cycling of Motoma LFP was 1485 mAh see Fig.3. After 500 cycles, capacity faded to 1375 mAh. Capacity fade between first and last cycle was 7.5%. Motoma LFP capacity during 1 C cycling was immutable for 175 cycles. Capacity started to fade since 175 cycles in an almost linear trend. Coulombic efficiency was around 100% throughout the cycling period.

Discharge curve of Panasonic NCR before aging is shown in Figure 4. Capacity drop was gradual for all used discharge currents. It was found that the capacity when discharging with a current of 0.1 C is 3290 mAh and at 0.2 C is 3240 mAh. Discharge capacity for 0.5 C and 1 C is identically 3180 mAh. Voltage is slowly decreasing over time for all discharge currents. Most significant drop of potential could be seen for 1 C.

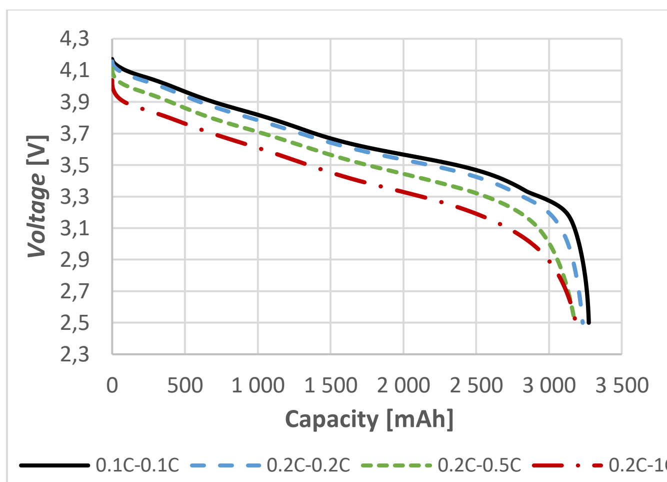


Figure 4 Discharging characteristics of Panasonic NCR

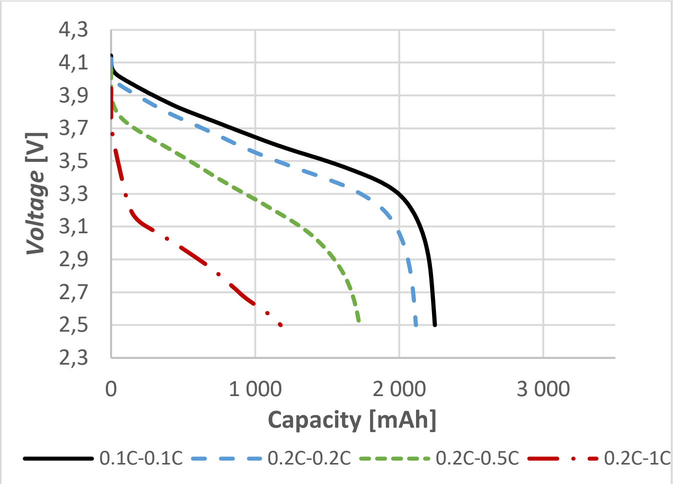


Figure 5 Discharging characteristics of aged Panasonic NCR

Capacity fade of Panasonic NCR after 500 cycles is most significant for higher C rates. Higher the C-rate, higher the capacity fade. Lowest 0.1 C rate discharge capacity is 2200 mAh, which is almost double to capacity 1180 mAh measured by 1C discharge. Total capacity fade is 33% for 0.1C discharge and 64% fade for 1 C discharge.

Capacity fade of Panasonic NCR is most significant up to 150 cycles. Cell lost 59% of its initial capacity. Since 150th cycle, capacity loss has become more stable. The cell lost another 25% of initial capacity over 350 cycles. Total capacity drop is nearly 85% after 500 cycles.

In contrast of Motoma LFP, the degradation of the Panasonic NCR was more noticeable. The capacity drop after 500 cycles is 85%, as shown in Figure 6. The discharge characteristics before and after cycling can be seen in Figures 4 and 5. The capacity even at a low load of 0.1 C is 30% lower. At 1 C load, the capacity already drops to 30% of the original capacity.

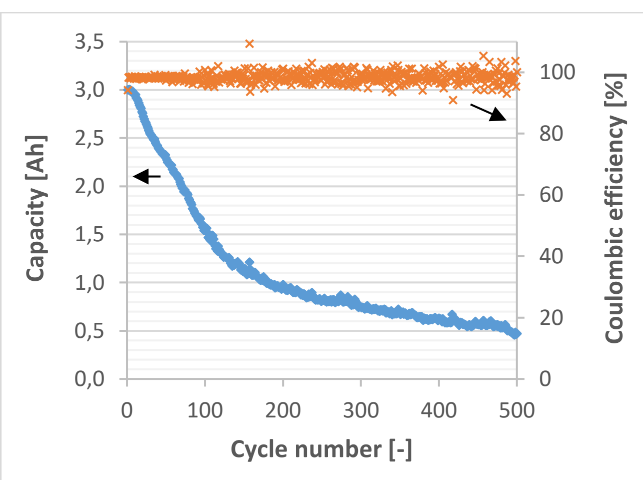


Figure 6 Capacity fade of Panasonic NCR during cycling

III. Cell disassembly

The aged cells were first discharged below 1V, so that there is not a significant amount of energy in the battery and there is no risk of fire in the event of a short circuit during heating. Subsequently, each battery was disassembled in the hood. First, the entire cell was weighed and then all the parts. The bill of materials of Motoma LFP and Panasonic NCR are shown in Figure 7 and Figure 8.

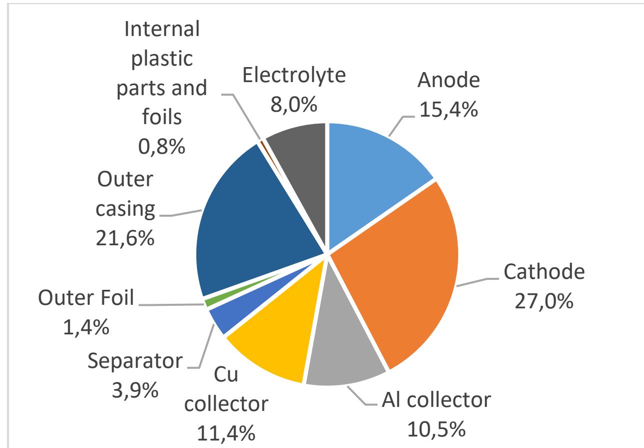


Figure 7 Bill of material of Motoma LFP cell

The electrode materials are the most significant parts, for Motoma LFP, cathode material is represented by 37.5%, of which 10.5% is aluminium collector. The anode material is represented by 26.4%, of which 11.4% is the copper collector. 21.6% of the total cell weight is the outer metal case. Total aluminium content in Motoma LFP cell is nearly 32%. Plastic parts represent about 2.2% of total weight. Separator content in cell is 3.9%. Total plastic content in Motoma LFP is 6.1%. Since it is not possible to determine the exact proportion of the electrolyte and its salts, content of electrolyte was determined this by the difference between the weight of the whole battery and the weight of all components after drying. Electrolyte content was determined to be 8%.

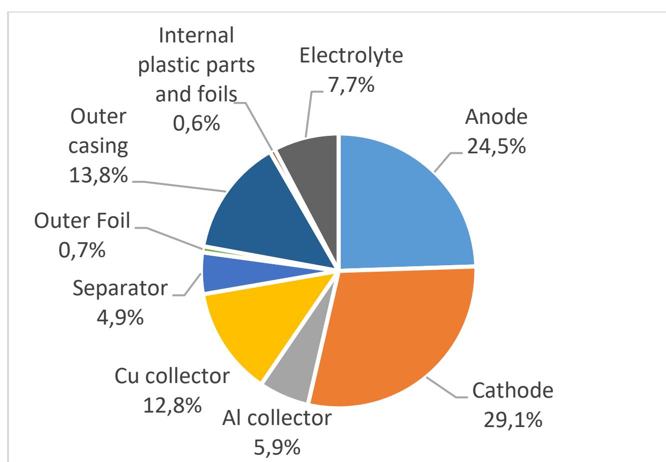


Figure 8 Bill of material of Panasonic NCR cell

The weight distribution of Panasonic NCR is 35% cathode material and aluminium collector. The NCA cathode material itself is represented by 29.1%. Aluminium collector makes up to 5.9% of cell weight. Outer aluminium case content is 13.8%. Total aluminium content in Panasonic NCR cell is 19.7%. Plastic parts and foils make up 1.3%, plastic separator 4.9%. Total plastic content of Panasonic NCR is 6.2%. The anode material is represented by 37.2%, of which 12.8% is copper collector. The electrolyte was calculated to be 7.7% of the cell weight.

For future analysis, anode material from each Motoma LFP cell and Panasonic NCR cell were separated from Cu collector by chemical solvents. Nine different solvents were used for this extraction. N-methyl-2-pyrrolidone (NMP), Toluene, N,N-dimethylacetamide (DMAC), N,N-dimethylformamide (DMF), Ethyl alcohol (EtOH), dimethylsulfoxide (DMSO), ethyl acetate (EtAc), dihydrolevoglucosenone (Cyrene) and Water. Anode was cut to proximately identical size and immersed into solvents. Several temperature were used for dissolving the cathode, from room temperature through 50, 75, 100 up to 120°C. Samples were observed for 24 hours total (after 1, 2, 5, 10, and 24 hours). Every used solvent volume was 30 ml.

Lithium plating was observed on the anode material of Motoma LFP cell. No other issues were observed for this battery cell. Lithium plating is shown in Figure 9.

No lithium plating was observed for Panasonic NCR cell. However, material was fragile and fell off the copper collector spontaneously. This flaking can be caused by prolonged heavy loading, where graphite undergoes changes, and these rapid changes can crack it. Anode material of Panasonic NCR is shown in Figure 10.

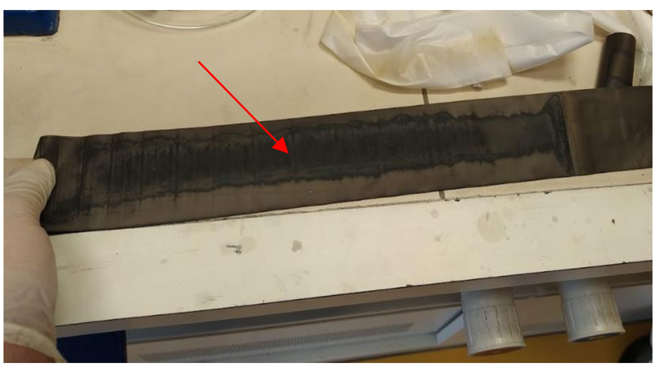


Figure 9 Lithium plating on the anode of Motoma LFP cell

Figure 10 Anode material of Panasonic NCR

Solvents like DMSO, DMF, DMAC and NMP are typical solvents for many cathode materials. [4] [5] However, the efficiency of these solvents for the anode material turned out to be practically zero. Toluene, EtOH and EtAc are cheap and available solvents. Nevertheless, from listed solvents, only water was capable of dissolving graphite material from copper current collector. Even at higher temperatures, solvents with close solubility parameters could not dissolve the sample. Dissolution in water was observed at room temperature in less than one hour as shown in Table 1. With increased temperature, dissolution was enhanced. Separated graphite material from current collector is shown in Figure 11.

As for the Motoma LFP, the Panasonic NCR had no reaction with the solvents used. The only solvent capable of separating the graphite and copper collector was water, as with the Motoma LFP. The water also dissolved the binder used for the Panasonic NCR at room temperature. The solvent efficiency for the Panasonic NCR is shown in Table 2.

Table 1 Solvent efficiency at 22°C for Motoma LFP anode dissolution

	22°C				
Solvent	1h	2h	5h	10h	24h
DMSO	x	x	x	x	x
DMF	x	x	x	x	x
DMAC	x	x	x	x	x
EtOH	x	x	x	x	x
H_2O	✓	-	-	-	-
Cyrene	x	x	x	x	x
EtAc	x	x	x	x	x
Toluene	x	x	x	x	x
NMP	x	x	x	x	x

Table 2 Solvent efficiency at 22°C for Panasonic NCR anode dissolution

	22°C				
Solvent	1h	2h	5h	10h	24h
DMSO	x	x	x	x	x
DMF	x	x	x	x	x
DMAC	x	x	x	x	x
EtOH	x	x	x	x	x
H_2O	✓	-	-	-	-
Cyrene	x	x	x	x	x
EtAc	x	x	x	x	x
Toluene	x	x	x	x	x
NMP	x	x	x	x	x

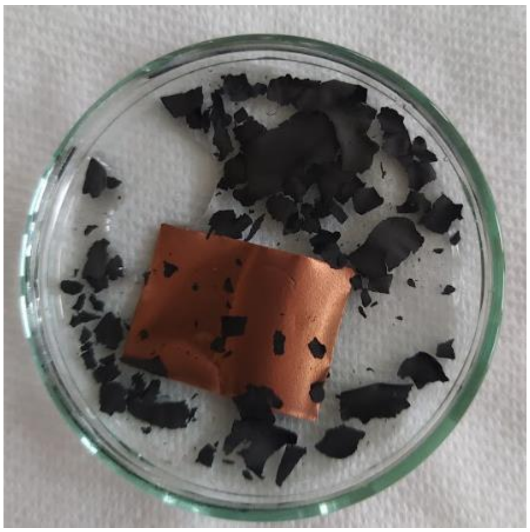


Figure 11 Dissolved cathode material of Motoma LFP

The graphite material of the anode of the Motoma LFP cell was separated in water from the copper collector in 100% of the area without the use of mechanical separation or other methods such as ultrasound.

IV. CONCLUSION

Direct recycling is an environmentally friendly option. Water as a solvent suitable for the anode material will reduce the ecological impact, which is considerable with pyrometallurgical, hence hydrometallurgical methods. As dissolution can be carried out at room temperature, energy consumption and carbon dioxide emissions are significantly reduced.

ACKNOWLEDGMENT

This work was supported by the specific graduate research of the Brno University of Technology No. FEKT-S-23-8286

BIBLIOGRAPHY

[1] F. Larouche, F. Tedjar, K. Amouzegar, G. Houlachi, P. Bouchard, G. Demopoulos and K. Zaghib, "Progress and Status of Hydrometallurgical and Direct Recycling of Li-Ion Batteries and Beyond", *Materials*, vol. 13, no. 3, 2020.

[2] Y. Miao, L. Liu, Y. Zhang, Q. Tan and J. Li, "An overview of global power lithium-ion batteries and associated critical metal recycling", *Journal of Hazardous Materials*, vol. 425, 2022.

[3] J. Edge, S. O'Kane, R. Prosser, N. Kirkaldy, A. Patel, A. Hales, A. Ghosh, W. Ai, J. Chen, J. Yang, S. Li, M. Pang, L. Bravo Diaz, A. Tomaszewska, M. Marzook, K. Radhakrishnan, H. Wang, Y. Patel, B. Wu and G. Offer, "Lithium ion battery degradation: what you need to know", *Physical Chemistry Chemical Physics*, vol. 23, no. 14, pp. 8200-8221, 2021.

[4] J. Wu, M. Zheng, T. Liu, Y. Wang, Y. Liu, J. Nai, L. Zhang, S. Zhang and X. Tao, "Direct recovery: A sustainable recycling technology for spent lithium-ion battery", *Energy Storage Materials*, vol. 54, pp. 120-134, 2023.

[5] Y. Lu, K. Peng and L. Zhang, "Sustainable Recycling of Electrode Materials in Spent Li-Ion Batteries through Direct Regeneration Processes", *ACS ES&T Engineering*, vol. 2, no. 4, pp. 586-605, 2022.

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

2D Physical Modelling of Semiconductor Equations for Verification of 1D Lumped-Charge Model of Bipolar Power Devices

1st Jan Miklas
Department of Power Electrical and Electronic Engineering
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
jan.miklas@vut.cz

2nd Petr Prochazka
Department of Power Electrical and Electronic Engineering
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
prochazkap@vut.cz

***Abstract*—The paper demonstrates a finite-element method (FEM) simulation model of semiconductor devices operation. Classical semiconductor equations employing drift, diffusion and generation-recombination transport of charge carriers are established and variational forms for FEM assembly are derived in detail, including a basic voltage and current boundary conditions. Mathematically, a system of mutually coupled nonlinear equations need to be solved, which requires extensive use of nonlinear iteration solver. The derived model was coded in Python by strict use of open source tools, mainly grouped around FEniCSx project. Graphic output figures and characteristics for basic semiconductor structures are presented to demonstrate the functionality of model.**

An ultimate goal of presented effort is derivation and verification of simplified semi-analytical model of Insulated-gate bipolar transistor (IGBT), including precise transient behavior.

This kind of model should be calibrated by use of measurement-obtained data, so the qualitative behavior of device physics at reasonable computational cost is of primary interest of presented FEM model as opposed to commercial device development tools aiming at precise quantitative outputs; which need to be experimentally calibrated even so.

As an additional step to simplified one-dimensional model usability verification, results of unusual way of experimental estimation of minority-carrier excess charge within power bipolar transistor collector and base during on-state is presented and compared to simulation result.

***Index Terms*—Power BTJ model, power BJT switching, IGBT transient model, transistor switching measurement, semiconductor device simulation, finite element method, lumped charge model**

I. Introduction

A. Semiconductor Equations

Let's establish the basic semiconductor equations governing the main phenomena in semiconductor devices, as described numerous times in literature ([1], [2], [3], [4]), following the notation listed below:

Notation:

ε	Permittivity of material (silicon)
ψ	Electric potential
$\mathbf{E}$	Electric field intensity $\mathbf{E} = -\nabla\psi$
q	Elementary charge
p, n	Holes and electrons concentration respectively
n_i	Intrinsic carrier concentration for given material
$N_\mathrm{A}, N_\mathrm{D}$	Acceptors and donors concentration[1]
$\mathbf{J}_p, \mathbf{J}_n$	Hole and Electron current pre unit area
μ_p, μ_n	hole and electron mobility
D_p, D_n	Diffusion constants (Fick's Law)
R_p, R_n	Recombination-generation rate of holes and electrons
τ_p, τ_n	hole and electron recombination lifetime
k	Boltzmann's constant
T	Thermodynamic temperature
$\frac{kT}{q}$	Thermal Voltage ($25,9\,\mathrm{mV}$ at room temperature)
$\mathbf{n}$	Facet normal pointing in direction out of the boundary

Poisson's Equation (*Gauss's Law*):

$$\nabla \cdot (\varepsilon \nabla \psi) = -q(p - n + N_D - N_A) \tag{1}$$

Carrier transport - drift-diffusion equations:

$$\mathbf{J}_p = \overbrace{qp\mu_p\mathbf{E}}^{\text{drift}} - \overbrace{qD_p\nabla p}^{\text{diffusion}} \tag{2}$$

$$\mathbf{J}_n = qn\mu_n\mathbf{E} + qD_n\nabla n \tag{3}$$

Continuity equations for hole and electron current:

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\nabla \cdot \mathbf{J}_p - R \tag{4}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot \mathbf{J}_n - R \tag{5}$$

[1]For simplicity, all dopant atoms are considered fully ionised.

DOI 10.13164/eeict.2023.181

The drift current - as embodied in any conductor - is driven by electric field, with velocity given by mobility of charged particle, while diffusion current is driven by carrier concentration gradient, as given by *Fick's Law* of any kind of diffusion.

The Recombination-generation term R is dependent on holes and electrons concentration and doping profile (impurity atoms concentration) and should not be omitted as it affects the PN-junction conditions significantly, though there is no one generally valid yet simple model. Carrier recombination originates in number of different physical mechanisms. The main ones like trap-assisted "Shockley-Read-Hall" (SRH), "Auger", band-to-band and surface recombination effects are not negligible. The number of independent variables in R increases with particular model complexity.

Typically a Shockley-Read-Hall recombination (SRH) model is used if no special effects are needed to be accounted, expressed by form [3], [4]:

$$R = \frac{n \cdot p - n_i^2}{\tau_p(n+n_0) + \tau_n(p+p_0)} \tag{6}$$

where n_i stands for intrinsic carrier concentration in given material, p_0, n_0 for thermal equilibrium concentrations of holes and electrons and τ_p, τ_n for holes and electron recombination lifetimes respectively.

Considering $\mathbf{E} = -\nabla\psi$, equations (1) - (4) can be rearranged into 3-equation system with unknowns (ψ, p, n) as follows:

$$\nabla \cdot (\varepsilon \nabla \psi) = -q(p - n + N_D - N_A) \tag{7}$$

$$\frac{\partial p}{\partial t} = \nabla \cdot (p\mu_p \nabla\psi + D_p \nabla p) - R \tag{8}$$

$$\frac{\partial n}{\partial t} = \nabla \cdot (-n\mu_n \nabla\psi + D_n \nabla n) - R \tag{9}$$

Equations (7) - (9) present a system of mutually coupled nonlinear partial differential equations, that can be discretized, assembled and solved numerically by finite elements method.

II. FEM Model

For purpose of this paper only a static semiconductor equations (steady state) need to be solved, i.e. the left-hand side of (4) and (5) equals zero.

To solve the system of coupled equations a mixed FEM needs to be utilized, where the unknown variables are treated as components of one possibly multidimensional "vector" variable. Using the mixed function space allows avoiding the manual and computationally expensive decoupling of equations by any kind of splitting scheme.

A. Scaling the Equations

Since the set of variables (ψ, p, n) as well as physical constants exhibit values difference of many orders of magnitude, it is advantageous to scale the equations [5] and geometric domain so that the problem becomes dimensionless. The values can be easily scaled back to physical dimensions after the solution by multiplying the dimensionless value by its scaling constant according to:

$$X = \frac{X_{\text{orig}}}{X_0} \tag{10}$$

where X stands for the scaled dimensionless value, X_{orig} for dimensional physical value and X_0 is the scaling factor.

For compact notation, all variables and values used in equations are considered scaled in the remainder of this paper unless indexed explicitly as "$_{\text{orig}}$". The differential operators also need to be scaled accordingly. The same applies for time in transient case solution.

Consider the original Poisson's equation (7) with dimension-emphasizing notation:

$$\nabla_{\text{orig}} \cdot (\varepsilon \nabla_{\text{orig}} \psi_{\text{orig}}) = -q(p_{\text{orig}} - n_{\text{orig}} + N_{\text{D,orig}} - N_{\text{A,orig}}) \tag{11}$$

After utilization of scaled quantities

$$\begin{aligned}
\psi_{\text{orig}} &= \psi \cdot \psi_0 \\
p_{\text{orig}} &= p \cdot p_0 \\
n_{\text{orig}} &= n \cdot n_0 \\
N_{\text{A,orig}} &= N_{\text{A}} \cdot N_0 \\
N_{\text{D,orig}} &= N_{\text{D}} \cdot N_0 \\
x_{\text{orig}} &= x \cdot X_0 \\
\frac{\mathrm{d}}{\mathrm{d}x_{\text{orig}}} &= \frac{\mathrm{d}}{\mathrm{d}x \cdot X_0} = \frac{1}{X_0}\frac{\mathrm{d}}{\mathrm{d}x} \\
\nabla_{\text{orig}} &= \frac{1}{X_0}\nabla
\end{aligned} \tag{12}$$

we get a scaled equation:

$$\lambda_0 \nabla \cdot (\nabla\psi) = -(p - n + N_D - N_A) \tag{13}$$

where

$$\lambda_0 = \frac{\varepsilon\psi_0}{X_0^2 q N_0} \tag{14}$$

Similarly the continuity equations (4),(5) at steady state ($\frac{\mathrm{d}p}{\mathrm{d}t} = 0, \frac{\mathrm{d}n}{\mathrm{d}t} = 0$) become:

$$0 = \lambda_1 \nabla \cdot (p\mu_p \nabla\psi + D_p \nabla p) - \lambda_1 R \tag{15}$$

$$0 = \lambda_1 \nabla \cdot (-n\mu_n \nabla\psi + D_n \nabla n) - \lambda_1 R \tag{16}$$

with

$$\lambda_1 = \frac{N_0 \psi_0 \mu_0}{X_0^2} \tag{17}$$

and

$$\begin{aligned}
\mu_{\text{p,orig}} &= \mu_{\text{p}} \cdot \mu_0 \\
\mu_{\text{n,orig}} &= \mu_{\text{n}} \cdot \mu_0 \\
D_{\text{p,orig}} &= D_{\text{p}} \cdot D_0, \text{ where } D_0 = \frac{kT}{q}\mu_0 \\
D_{\text{n,orig}} &= D_{\text{n}} \cdot D_0
\end{aligned} \tag{18}$$

The equation $\frac{D}{\mu} = \frac{kT}{q}$ used in (18) is known as *Einstein's relation* known from kinetic theory.

ψ_0	$\frac{kT}{q}$
X_0	$\max(x_{\mathrm{domain}})$
N_0	$\max(N_\mathrm{A}, N_\mathrm{D})$
μ_0	$\max(\mu_p, \mu_n)$
D_0	$\frac{kT}{q}\mu_0$

TABLE I
SCALING FACTORS.

The basic approach in order to maintain the scaled variables roughly close to unity is to normalize each quantity to its "typical" value, though it is strictly not possible (for example, p, n, N_A, N_D typically exhibit different values of at least an order of magnitude). The set of scaling factors used in presented model, partly inspired by [6], [4], [7] and [3], is summarized in Tab. I.

B. Problem Statement

Let us state the *boundary-value problem* according to (13), (15), (16) within a spatial domain Ω:

$$\lambda_0 \nabla \cdot (\nabla \psi) = -(p - n + N_D - N_A) \quad \text{in } \Omega \tag{19}$$

$$0 = \lambda_1 \nabla \cdot (p\mu_p \nabla \psi + D_p \nabla p) - \lambda_1 R(p,n) \quad \text{in } \Omega \tag{20}$$

$$0 = \lambda_1 \nabla \cdot (-n\mu_n \nabla \psi + D_n \nabla n) - \lambda_1 R(p,n) \quad \text{in } \Omega \tag{21}$$

with boundary conditions

$$\psi = \psi_\mathrm{BC} \quad \text{on } \Gamma_\mathrm{D0} \tag{22}$$

$$p = p_\mathrm{BC} \quad \text{on } \Gamma_\mathrm{D1} \tag{23}$$

$$n = n_\mathrm{BC} \quad \text{on } \Gamma_\mathrm{D2} \tag{24}$$

$$\mathbf{n} \cdot \nabla \psi = g \quad \text{on } \Gamma_\mathrm{N0} \tag{25}$$

Here, ψ, p, n are unknown functions, $N_\mathrm{D}, N_\mathrm{A}, \mu_\mathrm{p}, \mu_\mathrm{n}, D_\mathrm{p}, D_\mathrm{n}$ are normalized (scaled) material constants, λ_0, λ_1 the scaling factors and $R(p,n)$ the source term as defined by (6). Ω represents the spatial domain (device geometry) composed by subdomains with diverse material properties matching the various doping regions in device; and $\partial\Omega = \Gamma_\mathrm{D0} \cup \Gamma_\mathrm{D1} \cup \Gamma_\mathrm{D2} \cup \Gamma_\mathrm{N0}$ is the domain boundary emphasizing the Dirichlet and Neumann type of individual boundaries.

C. Variational Forms

The unknown variables are approximated by set of piecewise *trial functions* (ψ, p, n). System of equations (19) - (21) represent a strong form, i.e. the continuous *differential* form of equations with exact solution. To be able to get an approximate piecewise solution by FEM, the problem must be formulated into weak form by multiplying by corresponding *test functions* (v_0, v_1, v_2) - sometimes called the weight functions - and integrating over the whole domain. Moreover, the continuous *trial* and *test* function spaces must be replaced by discrete function spaces to enable the numerical approximate solution.

Doing so, the mixed function-space variational problem can be phrased as follows: find $(\psi, p, n) \in V_0 \times V_1 \times V_2$ such that

$$a\left((\psi, p, n), (v_0, v_1, v_2)\right) = L\left(v_0, v_1, v_2\right) \tag{26}$$

for all $(v_0, v_1, v_2) \in \hat{V}_0 \times \hat{V}_1 \times \hat{V}_2$; where

$$\begin{aligned} a\left((\psi, p, n), (v_0, v_1, v_2)\right) = \\ &- \int_\Omega \lambda_0 \nabla \psi \cdot \nabla v_0 \, \mathrm{d}x \\ &- \int_\Omega D_p \nabla p \cdot \nabla v_1 \, \mathrm{d}x - \int_\Omega \mu_\mathrm{p} p \nabla \psi \cdot \nabla v_1 \, \mathrm{d}x \\ &- \int_\Omega D_n \nabla n \cdot \nabla v_2 \, \mathrm{d}x + \int_\Omega \mu_\mathrm{n} n \nabla \psi \cdot \nabla v_2 \, \mathrm{d}x \end{aligned} \tag{27}$$

$$\begin{aligned} L(v_0, v_1, v_2) = \\ &- \int_\Omega (p - n + N_\mathrm{D} - N_\mathrm{A}) v_0 \, \mathrm{d}x - \int_{\Gamma_{N0}} g v_0 \, \mathrm{d}s \\ &+ R v_1 \, \mathrm{d}x \\ &+ R v_2 \, \mathrm{d}x \end{aligned} \tag{28}$$

and $\hat{V}_0 \times \hat{V}_1 \times \hat{V}_2 \subset V_0 \times V_1 \times V_2$ form a discrete *mixed function space* for test and trial (to be approximated) functions. Each of the subspaces represents a classic Lagrange "P_1" finite element.

The differential elements $\mathrm{d}x$ and $\mathrm{d}s$ denote integration over the domain Ω and over the domain boundary $\partial\Omega$ respectively. That means the surface integral over Ω and line integral over $\partial\Omega$ in the two-dimensional geometry.

D. Boundary Conditions

Among others, the two basic types of boundary conditions appear in semiconductor simulations literature: so called *voltage* and *current* boundary conditions to emphasize which of the external circuit value is forced to the device. Particular boundary conditions for semiconductor equations variables (ψ, p, n) need to be treated accordingly along with the type of contact that interfaces the device with external circuit - ohmic (metal), Schottky or insulator contact [4], [3] being the basal ones.

1) Voltage Boundary Conditions: The Dirichlet boundary conditions (22), (23), (24) constitute the basic form of ohmic-contact *voltage-type* boundary conditions. The value ψ_BC equals simply the value of electric potential applied to given contact (minus the *built-in potential* [8] which corresponds to equilibrium potential across PN-junctions and is compensated by semiconductor-metal contact potential in physical reality when device terminals are shorted). For $p_\mathrm{BC}, n_\mathrm{BC}$ values, several strategies have been published; as a very basic one, the equilibrium concentration values can be used, that assumes the boundaries to be distant enough from PN-junction depletion region so that all excess charge caused by forward bias can be considered to recombine, which is valid up until significant current or unusual geometries.

2) Current Boundary Conditions: Neumann boundary condition (25) or more sophisticated one must be derived for forced-current contacts.

Total current in any position equals:

$$\mathbf{J} = \mathbf{J}_p + \mathbf{J}_n \tag{29}$$

Substitution of (2), (3) for $\mathbf{J}_p$ and $\mathbf{J}_n$ yields

$$\frac{\mathbf{J}}{q} = \nabla\psi\left(-\mu_\mathrm{p} p - \mu_\mathrm{n} n\right) - D_\mathrm{p}\nabla p + D_\mathrm{n}\nabla n \tag{30}$$

From (30) we get the following expression for normal component of electric potential gradient as a Neumann condition, and we call it g:

$$g = \mathbf{n}\cdot\nabla\psi = \frac{D_\mathrm{n}\nabla n\cdot\mathbf{n} - D_\mathrm{p}\nabla p\cdot\mathbf{n} - \frac{\mathbf{J}}{q}\cdot\mathbf{n}}{\mu_\mathrm{n} n + \mu_\mathrm{p} p} \tag{31}$$

It may be worth noting p and n represent scalar fields so that ∇p, ∇n and $\mathbf{J}$ are all vector quantities and the dot product with normal vector $\mathbf{n}$ produces scalar values.

All of the other boundaries are treated as natural "reflective" boundaries with zero gradient of quantities in normal direction.

E. Solver and Initial Conditions

Since the system is nonlinear, the bilinear form (26) should be reformulated into semilinear form F

$$\begin{gathered} F\left((\psi, p, n); (v_0, v_1, v_2)\right) = 0 \\ \forall (v_0, v_1, v_2) \in \hat{V}_0 \times \hat{V}_1 \times \hat{V}_2 \end{gathered} \tag{32}$$

such that

$$F = a - L = 0 \tag{33}$$

with no need to treat the unknowns (ψ, p, n) specially as trial functions in direct method; an appropriate iteration solver needs to be employed instead - most commonly the Newton's method.

Dividing by function derivative in a standard single variable Newton's method must be replaced by multiplication by system's inverse $k \times k$ *Jacobian* matrix.

Setting the initial conditions consists of assigning certain values to unknown functions ψ, p, n prior to performing Newton's iterations. To make solver general, the initial guess values should be provided algorithmically, possibly violating the physical sense. Of course, consistency with Dirichlet boundary conditions are desirable whenever possible.

In proposed model, zero potential ($\psi = 0$) and majority charge carriers equal to doping profile ($p = N_\mathrm{A}, n = N_\mathrm{D}$) are used as initial guess. For situations greatly exceeding the equilibrium conditions, the stepping of bias voltage, current or even carrier concentrations is preformed to preserve convergence.

III. Implementation

The FEM model is coded in Python by use of open source tools, mainly grouped around FEniCSx project [9], [10]. It is comprised of the libraries UFL [11], Basix [12], [13], FFCx [14] and DOLFINx. All of the project components are accessible via high-level Python and C++ interfaces.

The variational forms are formulated in UFL language and DOLFINx provides most of the methods for defining mesh, tagging individual FEM cells, assigning material properties and boundary conditions according to cell tag, input/output file handling, direct methods for solving linear problem formulated in UFL and even automated method for computing of Jacobian matrix and evaluating of residua (with boundary conditions taken into account) via function `NonLinearProblem`.

For illustration, the variational forms defined by (26) can be transcribed into almost 1:1 Python/UFL code as follows:

```
# Poisson:
a0 = inner(grad(Psi), grad(v0)) * dx - 1/lmda0 * (Nd
    -Na-n+p + EPS) * v0 * dx # EPS to avoid zero
# Neumann BC on boundary tag_bc_B
g = (Dn*inner(grad(n), nn) - Dp*inner(grad(p), nn) -
     Jb) / (mob_n*n + mob_p*p)
L0 = - g * v0 * ds(tag_bc_B)

# Recombination / Generation rate
RG = (ni**2-n*p) / (Tau_p*(n+n_eq) + Tau_n*(p+p_eq))
# Continuity p:
a1 = -D_p * inner(grad(p), grad(v1)) * dx - mob_p *
    p * inner(grad(Psi), grad(v1)) * dx
L1 = RG *v1* dx
# Continuity n:
a2 = -D_n * inner(grad(n), grad(v2)) * dx + mob_n *
    n * inner(grad(Psi), grad(v2)) * dx
L2 = RG *v2* dx

F = a0+a1+a2 -L0-L1-L2
```

The domain, mesh, subdomain tags and boundary conditions tags for all cells are created in gmsh [15].

Post-processing and visualizations are made directly in python program by use of matplotlib library [16] using output data formatted by appropriate DOLFINx and numpy [17] functions.

IV. Demonstration - PN Junction Diode

To demonstrate the model functionality, an PN-diode of rectangular subdomains representing uniformly doped P-type and N-type regions was modeled. N-region is background

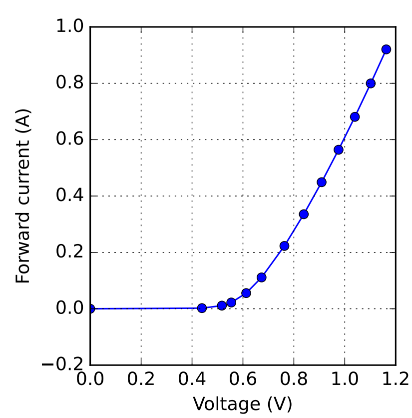


Fig. 1. Simulated static V-A characteristic of simple PN diode. Horizontal axis represents an external applied voltage i.e.with built-in voltage accounted. Current axis is normalized by arbitrary choice of device cross-section area.

doped by acceptors equal to P-region doping concentration, emulating the P-type base material in device manufacturing process.

A. Geometry Domain and Subdomains

Geometry and material properties of simulation subdomains is defined as follows:

- P-region:
  $N_\mathrm{D} = 0$, $N_\mathrm{A} = 1.3 \cdot 10^{16}\,\mathrm{cm}^{-3}$, length $1.25\,\mu\mathrm{m}$
- N-region:
  $N_\mathrm{D} = 1.6 \cdot 10^{16}\,\mathrm{cm}^{-3}$, $N_\mathrm{A} = 1.3 \cdot 10^{16}\,\mathrm{cm}^{-3}$, $1.25\,\mu\mathrm{m}$

The simulation results are scalable through cross-section area, so there is no need to define an exact cross-section dimension.

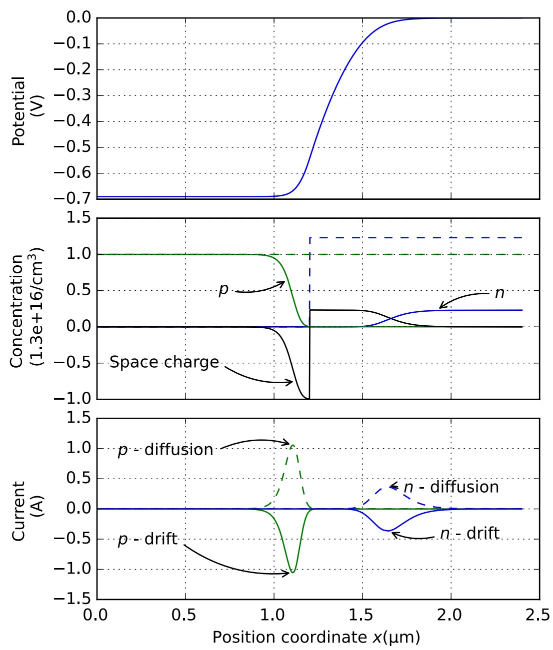


Fig. 2. PN junction electrostatics in thermal equilibrium. Doping concentrations N_D and N_A are drawn by blue and green dashed line.

B. Boundary Conditions

- P-side: *current* BC with value sweep,
- N-side: *voltage* BC set to $\psi_\mathrm{BC,N} = 0$ (cathode grounded)

C. Simulation Results

The thermal equilibrium results of PN-junction electrostatics, which are virtually the same as reverse bias (except the depletion region width and potential extent) due to existence of built-in potential, are depicted on Fig. 2. The depletion region free of mobile carriers - thus non-zero net charge $q(p-n+Nd-Na)$ - is formed at junction vicinity. A positive drift current is perfectly compensated by negative diffusion current and vice versa at any position.

In contrast, a forward biased PN-junction results are shown on Fig. 3. The depletion region is reduced greatly, minority carriers are injected from opposite regions and total current $\mathbf{J} = \mathbf{J}_p + \mathbf{J}_n$ is constantly flowing through device from P-anode to N-cathode.

A sequence of steady state situations like this with parametric sweep of total current can be arranged into current-voltage characteristic as shown on Fig. 1.

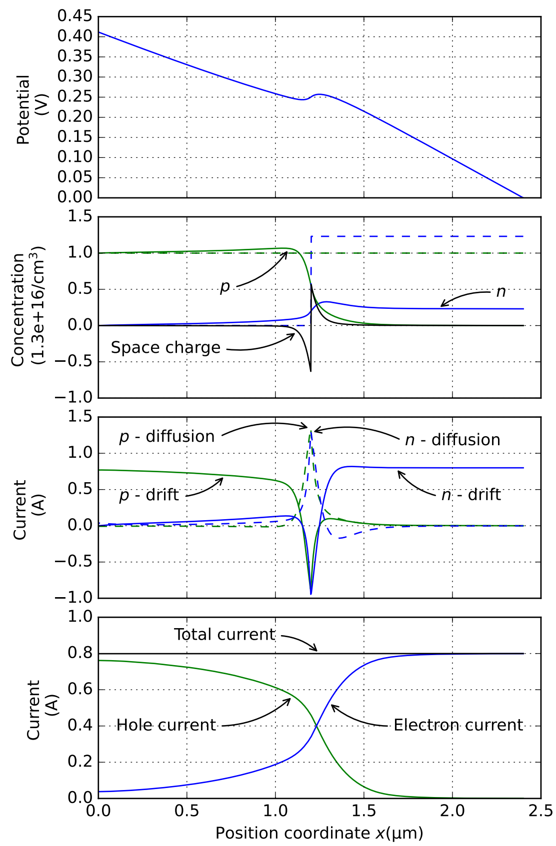


Fig. 3. PN junction electrostatics under forward bias condition.

V. NPνN Power Transistor - ON-State Excess Charge Storage

A. Background and 1D Model

NPνN nomenclature signifies a presence of very lightly doped so called N-drift region [18] in collector to support high voltage during reverse bias of base-collector junction. An example of such structure - representing one of many cells incorporated in real power transistor chip - is depicted on Fig. 4. Essentially the same structure is employed in any IGBT transistor. N-drift region has significant impact on transistor's static characteristic and slows-down switching as it can store a huge amount of excess charge.

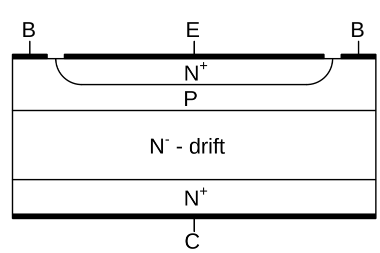


Fig. 4. NPνN power BJT domain geometry and boundaries definition.

Simplified analytical one-dimensional models, often called the lumped-charge models, that model an excess charge within device are widely used for explanation of power transistor operation [19], [20], being essentially an extension of traditional charge-control approach introduced by device modelling pioneers like Gummel [21].

A brief analysis 1D model of power BJT as shown on Fig. 5 is presented in our previous work [22]. The article introduces

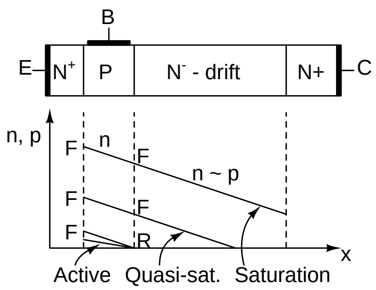


Fig. 5. One-dimensional model of power BJT excess charge storage at various operating stages. “F” and “R” stands for forward and reverse junction bias.

an experimental way of stored-charge estimation based on analysis of base drive current during turn-off transient. Due to inability to access the base current within IGBT, a power BJT was chosen for characterization to support the simplified model experimentally.

One of the outputs examining the stored charge under variable collector current and showing the transition between active and quasi-saturation mode of operation is presented on Fig. 6. The rationalisation of shape of the curves is discussed

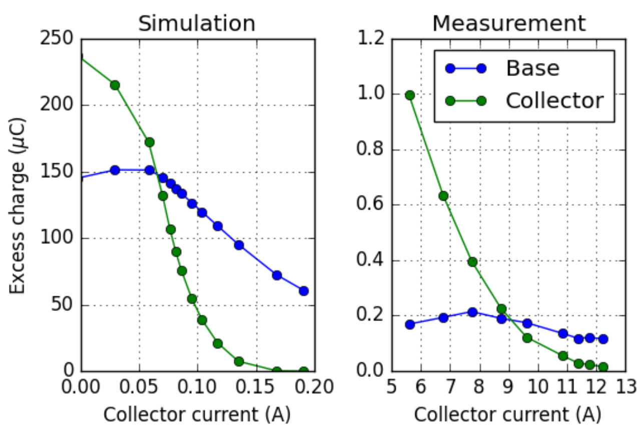


Fig. 6. Evaluation of stored charge within base and collector at variable collector current based on model and experiment [22].

in [22] and is not an objective of this paper. Here only a physical simulation support for the experiment is supplied.

B. 2D Simulation Model

Simulation domain adheres geometry, subdomains and boundaries depicted on Fig. 4 The subdomains are uniformly doped according to:

	N^+ Emitter	P Base	N^- - drift	N^+ Collector
Layer Width	$4\,\mu$m	$6\,\mu$m	$15\,\mu$m	$6\,\mu$m
N_D (cm^{-3})	$5 \cdot 10^{18}$	$1 \cdot 10^{15}$	$1 \cdot 10^{15}$	$1 \cdot 10^{18}$
N_D (cm^{-3})	$1 \cdot 10^{17}$	$1 \cdot 10^{17}$	0	0

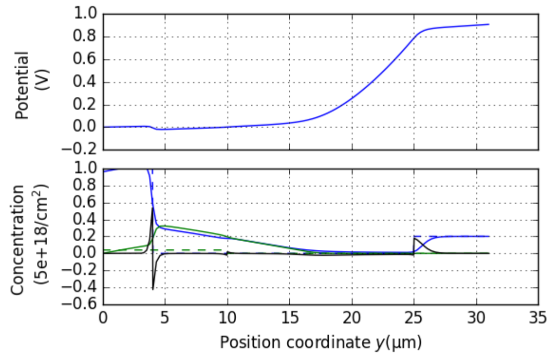


Fig. 7. Simulation showing excess charge storage in quasi-saturation mode. The line colors correspond with those described on Fig. 2.

Voltage boundary condition ψ_{BC} was applied on emitter contact. Parametric sweep of base and collector voltage boundary conditions were performed to achieve any of the operation modes as shown on current-voltage characteristic on Fig. 10. An example of solution of unknown functions (ψ, p, n) at quasi-saturation bias mode is shown on Fig. 8 and visualisation of vector-valued variables $\mathbf{J}, \mathbf{J}_p, \mathbf{J}_n$ are shown on Fig. 9 - the transistor effect causing current amplification is clearly apparent.

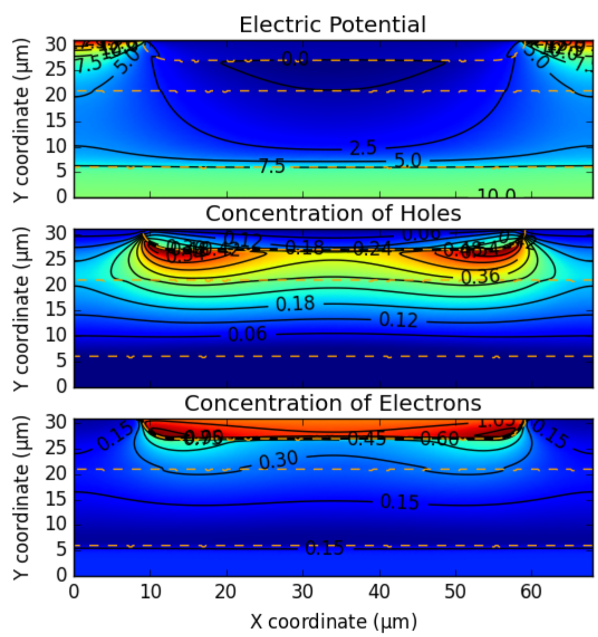


Fig. 8. Solution of Electrostatic variables (ψ, p, n) at quasi-saturation mode.

VI. Conclusion

The simulation outputs validate the feasibility of building an open-source implementation of semiconductor device physical model. It is demonstrated on simplified uniformly doped semiconductor structures, mainly for geometry and mesh coding

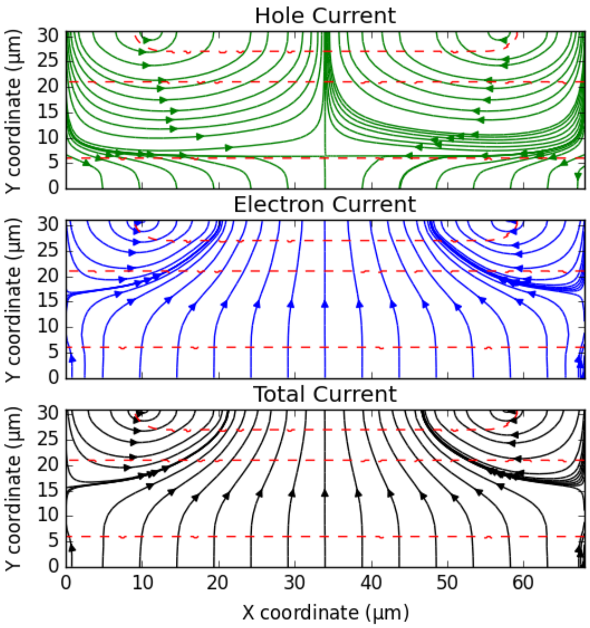


Fig. 9. Streamlines of currents in quasi-saturation mode.

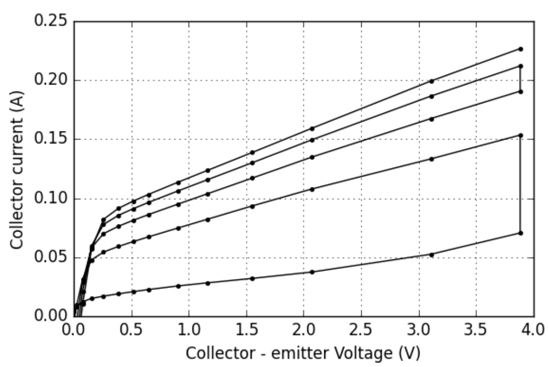


Fig. 10. Simulated output characteristic of NPνN transistor. The slope of the curves is given mainly by N-drift region resistivity in quasi-saturation mode.

simplicity; though it fully validates of numerical approach as well as qualitative explanation of device internal physics in sufficient detail.

The basic comparison with real measurement data of excess charge storage within power BJT shows good qualitative correlation with simulation data proving the FEM model to serve as good validation tool for analytical assumptions in further development of simplified models.

The main appeal in latter model version can be defined as:

- initial guess accuracy improvement for arbitrary bias conditions, being apparently one of the most crucial factors affecting the numerical convergence,
- more reasonable physics-based boundary conditions for any bias,
- extension of static transport equations to time dependent problem using either backwards Euler or any advanced scheme, allowing to model the transient behavior,
- more realistic non-uniform doping profiles and material properties definitions,
- advanced physical and/or empirical models of physical parameters (mobility, recombination, etc.).

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).

References

[1] R. Pierret, *Advanced Semiconductor Fundamentals*, ser. Modular series on solid state devices. Prentice Hall, 2003.
[2] S. Sze and K. Ng, *Physics of Semiconductor Devices*. Wiley, 2006.
[3] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Springer Vienna, 1984.
[4] M. R. Pinto, C. S. Rafferty, and R. W. Dutton, *PISCES II: Poisson and Continuity Equation Solver*. Stanford Electronics Laboratoories, Stanford University, Stanford Ca 94305, 1984.
[5] H. P. Langtangen and G. K. Pedersen, *Scaling of Differential Equations*. Springer, 2106. [Online]. Available: https://hplgit.github.io/scaling-book/doc/pub/book/pdf/scaling-book-4screen-sol.pdf
[6] A. De Mari, "An accurate numerical steady-state one-dimensional solution of the p-n junction," *Solid-State Electronics*, vol. 11, no. 1, pp. 33–58, 1968.
[7] Z. Yu, D. Chen, L. So, and R. Dutton, *PISCES-2ET and Its Application Subsystems*. Integrated Circuits Laboratory, Stanford University, Stanford, California 94305, 1994.
[8] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics, Vol. II: The New Millennium Edition: Mainly Electromagnetism and Matter*, ser. The Feynman Lectures on Physics. Basic Books, 2011.
[9] "Fenicsx | fenics project," https://fenicsproject.org, (accessed Mar. 12, 2023).
[10] "Fenics project · github," https://github.com/FEniCS, (accessed Mar. 12, 2023).
[11] M. S. Alnaes, A. Logg, K. B. Ølgaard, M. E. Rognes, and G. N. Wells, "Unified form language: A domain-specific language for weak formulations of partial differential equations," *ACM Transactions on Mathematical Software*, vol. 40, 2014.
[12] M. W. Scroggs, I. A. Baratta, C. N. Richardson, and G. N. Wells, "Basix: a runtime finite element basis evaluation library," *Journal of Open Source Software*, vol. 7, no. 73, p. 3982, 2022.
[13] M. W. Scroggs, J. S. Dokken, C. N. Richardson, and G. N. Wells, "Construction of arbitrary order finite element degree-of-freedom maps on polygonal and polyhedral cell meshes," *ACM Transactions on Mathematical Software*, 2022, to appear.
[14] R. C. Kirby and A. Logg, "A compiler for variational forms," *ACM Transactions on Mathematical Software*, vol. 32, 2006.
[15] Geuzaine, Christophe and Remacle, Jean-Francois, "Gmsh." [Online]. Available: http://http://gmsh.info/
[16] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
[17] C. R. Harris *et al.*, "Array programming with NumPy."
[18] B. Baliga, *Fundamentals of Power Semiconductor Devices*. Springer US, 2010.
[19] A. R. Hefner, "An improved understanding for the transient operation of the power insulated gate bipolar transistor (igbt)," *IEEE Transactions on Power Electronics*, vol. 5, no. 4, pp. 459–468, 1990.
[20] a. R. Hefner, "Modeling buffer layer igbts for circuit simulation," *IEEE Transactions on Power Electronics*, vol. 10, no. 2, pp. 111–123, 1995.
[21] H. K. Gummel, "A charge control relation for bipolar transistors," *The Bell System Technical Journal*, vol. 49, no. 1, pp. 115–120, 1970.
[22] J. Miklas and P. Prochazka, "Identification of power bjt operating stages based on experimental excess charge estimation," in *2021 IEEE 19th International Power Electronics and Motion Control Conference (PEMC)*, 2021, pp. 615–621.

Comparison of machine learning training sampling schemes for induction machine modeling

1st Vladimír Bílek
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
vladimir.bilek@vut.cz

***Abstract*—The aim of the paper is to demonstrate the modeling of an induction machine using a chosen machine learning technique, followed by a comparison of the training sampling schemes for this technique. A simple 3-phase induction machine with an axially slitted solid rotor has been selected for the case study, where FEM-based program Ansys Electronics Desktop has been used for its calculation. A total of 3 training schemes were considered and compared with each other for the machine learning technique. Some of the comparison results are given and discussed at the end of this paper. The described methodology can be used to accelerate the design and optimization of any type of electrical machine.**

***Index Terms*—FEA, Finite element method, Gaussian process regression, Induction machine, Machine learning, Solid rotor, Surrogate modeling**

I. Introduction

Nowadays, the use of numerical methods is fundamental for the design of electrical machines. Compared to measurements, they are able to calculate electromagnetic models with high accuracy. Finite element method (FEM) is the most widely used numerical method in the field of electrical machine design [1]. The optimization of electrical machines, however, involves very costly and time-consuming calculations [2]. Due to the ever-increasing power consumption and rising energy prices, it can be very financially costly, especially using large computing servers. Therefore, new ways of calculating and optimizing electrical machines that are much less time-consuming and costly are being sought. The best solution seems to be the use of machine learning and artificial intelligence, which is very popular in other technical fields.

In the following sections, a machine learning technique applied to a selected case study machine will be presented. The main focus of the paper will be to compare the training sampling schemes (TSS) for surrogate modeling and determine which one is the best for the design of electrical machines.

II. Used methodology

According to [3], surrogate modeling replaces expensive simulations or engineering problems. It is a special case of machine learning technique, replacing all expensive tasks with approximate functions, that are much faster to evaluate. Approximate functions (also called surrogate models or metamodels) are mathematical functions representing machine learning technique. Surrogate models are created from training data, which consists of a small number of expensive simulations calculated based on TSS. Using the approximated training data, it predicts the output of the original model. Due to approximation, surrogate models always exhibit some error between the predicted and the real data. It is always necessary to minimize this error by using proper machine learning technique and a suitable TSS. The biggest advantage of surrogate models is the ability to calculate large amounts of data (hundreds of thousands to millions) in a very short time. The workflow of surrogate modeling is shown in Fig. 1.

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.

A. Machine learning technique

There are several machine learning techniques for the assembly of surrogate models. Each technique is suitable for a different situation, so choosing the right one is crucial. Following [4], the machine learning techniques can be:

- polynomials,
- kriging,
- artificial neural networks (ANN), and
- radial basis function.

In this case, the Kriging was chosen since it is a very popular machine learning technique for surrogate modeling that is highly favored in terms of its flexibility, with the ability to model a wide range of functions while requiring only a handful of training data samples. This is possible due to the mathematical basis of the Gaussian process model (GPM).

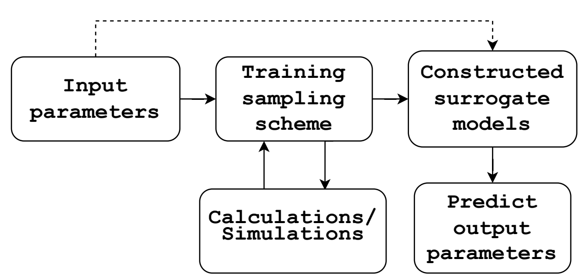


Fig. 1. Workflow of surrogate modeling.

GPM is a supervised machine learning method that is used to solve regression or classification problems [3]. The model has two parts: a deterministic polynomial term and the realization of a stochastic process that accounts for the lack of fit in the polynomial term. Based on the training data, this means the model predicts what the value of the original model might be based on a single input vector and what its overall global trend is. It also determines in which region of the modeled space it is least confident. This is illustrated on a simple 1D example in Fig. 2. The figure shows a real function (dashed line). For the training data, 6 observations (blue dots) are calculated for the GPM. The prediction of the surrogate model (red line) indicates the global trend of the predicted values. The orange fill indicates the uncertainty of the surrogate model prediction and is referred to as the 95% confidence interval. It is evident that the model has a high prediction error in the lowest confidence region.

For a GPM, it is necessary to define a kernel that provides the actual approximation of the data between two samples. There are several types of kernels, including periodic or non-periodic kernels, but, as stated in [5], the most popular kernel is the Radial Basis Function (RBF) kernel:

$$k(x_\mathrm{i}, x_\mathrm{j}) = \sigma_\mathrm{f}^2 \exp\left(-\frac{d(x_\mathrm{i}, x_\mathrm{j})^2}{2l^2}\right), \tag{1}$$

where $d(x_\mathrm{i}, x_\mathrm{j})$ is absolute difference between two samples x_i and x_j, σ_f are covariance scaling values, and l is a length-scale parameter ($l > 0$). The length-scale parameter l is a type of hyper-parameter whose value defines model uncertainty as well as the data prediction. This parameter must be optimized using the hyper-parameter optimization tuner before constructing a surrogate model. In the case of multiple input parameter dimensions, it is defined as a vector for each input parameter. However, the RBF kernel was found to exhibit poor performance on the data for electrical machines. Thus, another kernel was chosen, namely, the Matérn kernel [6]. The Matérn kernel is a generalization of the RBF kernel:

$$\begin{aligned} k(x_\mathrm{i}, x_\mathrm{j}) = \sigma_\mathrm{f}^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{\sqrt{2\nu}}{l} d(x_\mathrm{i}, x_\mathrm{j})\right)^{\nu} \cdot \\ \cdot K_\nu \left(\frac{\sqrt{2\nu}}{l} d(x_\mathrm{i}, x_\mathrm{j})\right), \end{aligned} \tag{2}$$

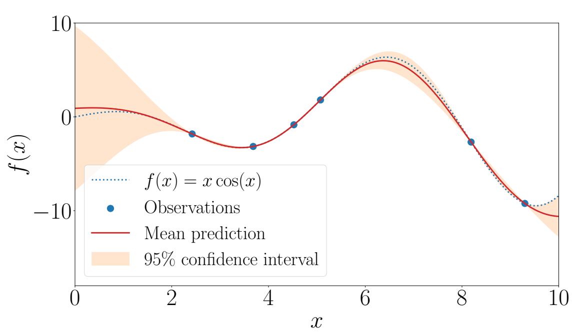


Fig. 2. Example of a Gaussian process regression for a 1D function.

where ν determines how smooth the resulting function is ($\nu > 0$). There are 4 smoothness values that are the most popular and used. In the case of $\nu = 1/2$ the Matérn kernel is identical to the exponential kernel. For $\nu = \infty$, the Matérn kernel becomes an RBF kernel. The RBF kernel is infinitely differentiable, which means that the approximated function is very smooth as well. Electrical machines often do not have smooth functions between input and output parameters. Therefore, for smoothness, once ($\nu = 3/2$) or twice ($\nu = 5/2$) differentiable is preferable to take into account the non-smooth functions. These 4 smoothness values are computed very quickly. For any different value, calculations can be very time-consuming and may not be very accurate. It was found that the best smoothness value with the highest accuracy is $\nu = 5/2$:

$$\begin{aligned} k(x_\mathrm{i}, x_\mathrm{j}) = \sigma_\mathrm{f}^2 \left(1 + \frac{\sqrt{5}}{l} d(x_\mathrm{i}, x_\mathrm{j}) + \frac{5}{3l} d(x_\mathrm{i}, x_\mathrm{j})^2\right) \cdot \\ \cdot \exp\left(-\frac{\sqrt{5}}{l} d(x_\mathrm{i}, x_\mathrm{j})\right) \end{aligned} \tag{3}$$

B. Training sampling schemes for the surrogate modeling

In surrogate modeling, defining the TSS is called the Design of Experiments. A total of 3 TSS were considered: General Full-Factorial (GFFS), Latin-Hypercube (LHS), and Box-Behenken (BBS). For GFFS [7] each model input parameter is given as a factor. The factors have a defined number of samples, which is, e.g., determined by a sensitivity analysis of the original model. All factors are multiplied among themselves, resulting in a final number of samples for the training data, which are uniformly distributed in all dimensions and space, forming a cubic structure Fig. 3(a). For LHS [8], the training data samples are scattered randomly Fig. 3(b), and their total number is defined uniformly. Several types of random sample generation exist, such as centering the samples in the middle of the modeled space, maximizing the minimum distance between samples, avoiding correlation between samples, etc. Finally, the BBS [9], which, like the GFFS, is uniformly distributed in a cubic structure, with the exception that the cube corners are empty without samples. Moreover, it is possible to define the number of samples in the middle of the modeled space Fig. 3(c).

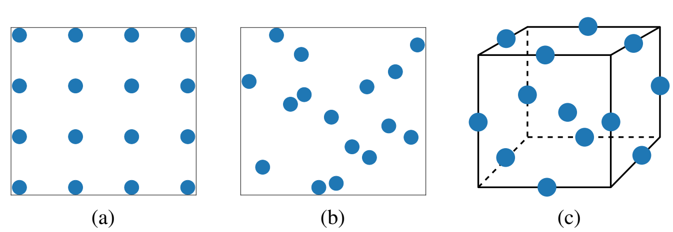


Fig. 3. Considered training sampling schemes for the surrogate modeling: General Full-Factorial (a), Latin-hypercube (b), and Box-Behenkan (c).

TABLE I
RATED ELECTROMAGNETIC MACHINE PARAMETERS

Parameter	Value
Output power (W)	589
Torque (Nm)	4.26
Input voltage (V)	3x400(Y)
Input current (A)	2.62
Frequency (Hz)	50
Speed (rpm)	1322
ElMag. efficiency (%)	68.32
Power factor (-)	0.437

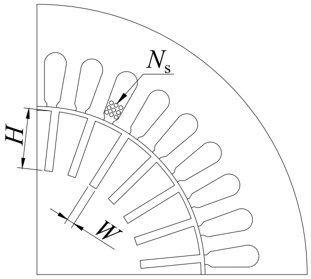


Fig. 4. Machine cross-section.

III. CASE STUDY

A. Analyzed machine

For the case study, a simple 3-phase, 4 pole induction machine (IM) with single-layer winding and an axially slitted solid rotor, was chosen. Rated electromagnetic parameters are given in Table I. The implemented materials are M470-50A for the stator and 1008 stainless steel for the rotor.

The machine input parameters selected for surrogate modeling were: active length, slit depth (H), slit width (W), and number of conductors (N_s). The monitored output parameters are: torque, electromagnetic efficiency, power factor, and torque ripple. By [4], it is not recommended to use more than 20 input parameters for GPM. Here, this condition is satisfied. A sketch of selected input parameters is illustrated in Fig. 4.

B. Process of surrogate modeling and deployment of surrogate models

Based on a sensitivity analysis of the case study machine, the sample size for TSS was estimated to be approximately 625. The paper [9] states that BBS creates a small number of samples for the TSS. For 4 input parameters, considering multiple points in the center of the modeled space, it would only be 27-29 samples. Therefore, BBS was excluded and only GFFS and LHS were considered for this study.

To process the initial training data, 3 programs were used. The first is Python, which automates the whole process of generating TSS using pyDOE [10], rescaling it for input data, and sending it to the electromagnetic model for subsequent electromagnetic calculations and storage of all data. The second is Ansys Electronics Desktop, which calculated the electromagnetic model using FEM. Due to the possibility of connecting Ansys and Python, it was feasible to perform up to 20 parallel calculations simultaneously. The electromagnetic model was set up for 2D transient analysis, wherein the model was set to constant speed to eliminate the additional variable of machine slip. With constant speed, it is easier to compare the monitored parameters between the calculated models. In 2D simulations, it is necessary to include 3D aspects of the rotor ends for solid rotor IM. This is done with the correction end-effect factor to include the rotor end resistance [11]. Here, to include the rotor end resistance, the conductivity of the rotor material is corrected using a simple equation:

$$\sigma_{\mathrm{Corr}} = \sigma_{\mathrm{Fe}} k, \tag{4}$$

where σ_{Fe} is the material conductivity and k is the corrective end-effect factor. The correction factor k is the product of two correction factors regarding machine geometry and slip. A modified Russell formula [11] was chosen to correct the material conductivity with respect to the machine geometry:

$$k_{\mathrm{Russell,M}} = 1 - \frac{\tau_{\mathrm{p}}}{\pi l_{\mathrm{s}}} \frac{\tanh\left(\frac{\pi l_{\mathrm{s}}}{\tau_{\mathrm{p}}}\right)}{\left(1 + \tanh\left(\frac{\pi l_{\mathrm{s}}}{\tau_{\mathrm{p}}}\right) \tanh\left(\frac{\pi l_{\mathrm{end}}}{\tau_{\mathrm{p}}}\right)\right)}, \tag{5}$$

where τ_{p} is rotor pole pitch, l_{s} is half of the machine's active length, and l_{end} is rotor end length. The machine slip correction factor was chosen according to Pyrhonen, Aho, and Nerg [11]:

$$k_{\mathrm{PAN}} = 1 - c\omega_{\mathrm{r}}^{\frac{3}{4}}, \tag{6}$$

where c is an experimental adaptation coefficient (estimated as $c = 0.022$) and w_{r} is angular slip frequency of the rotor.

As a final step, Excel stored the initial training data. For a small number of observations (625), an Excel or text file is sufficient. For hundreds of thousands to millions of observations, a database would be more appropriate. The process of obtaining initial training data is illustrated in Fig. 5.

With obtained training data, surrogate models were constructed. One model for each output parameter (4~in total). All training data were normalized (the maximum value is 1), to make it easier to approximate the data for parameters with similarly large ranges. As a typically recommended procedure, random shuffling and splitting of the normalized data into training and verification data followed. The verification data are used to validate the initial accuracy of the surrogate models. For smaller datasets (~10 000), it is recommended to split them into training and verification in the ratio of 90:10 to 80:20. For larger data sets (~1,000,000) this can be up to a ratio of 60:40 to 50:50. In this case, a 90:10 ratio was applied. The processed training data were used to construct surrogate models using GPM regression and Matérn kernel utilizing the SciKit-Learn [12]. To reproduce the results and even out conditions, both training data sets were shuffled and split based on the same seed.

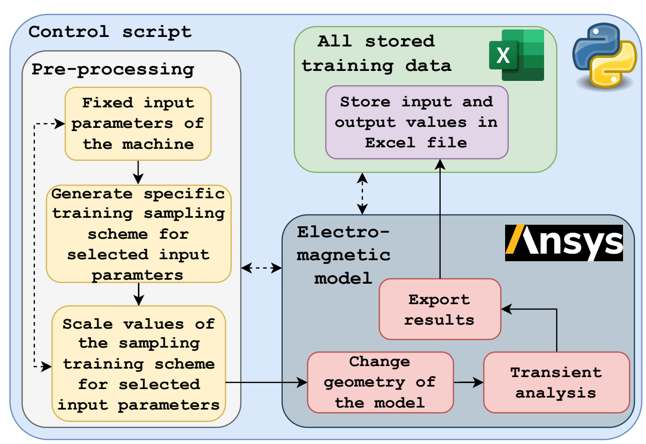


Fig. 5. Block diagram showing the algorithm for calculating the defined TSS.

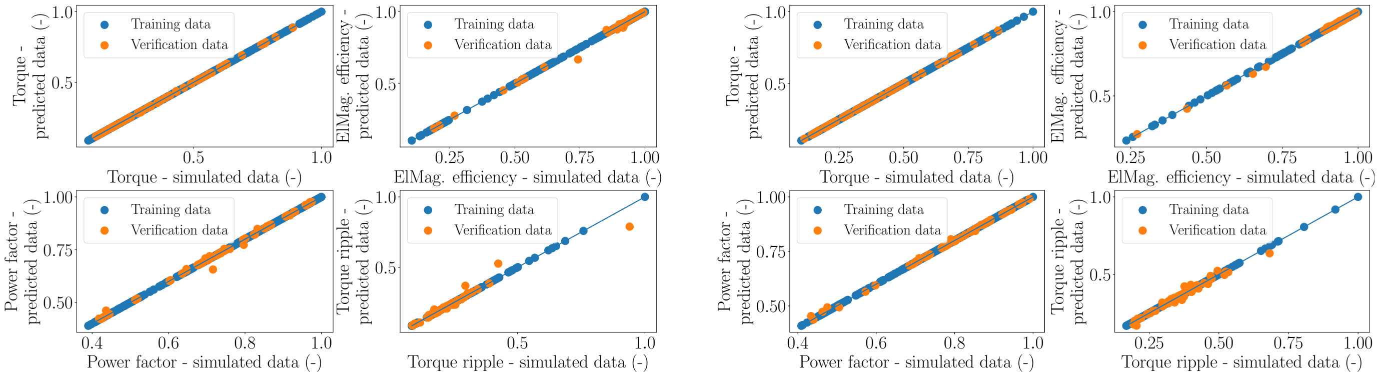


Fig. 6. Accuracy determination of the constructed surrogate models using GFFS (a) and LHS (b) TSS for: Torque, Electromagnetic Efficiency, Power Factor, and Torque Ripple.

C. Results and discussion

The accuracy of the surrogate model is determined by plotting the predicted (y-axis) and simulated (x-axis) training/verification data, as show in Fig. 6. Ideally, all points should be in one line. This mainly concerns training data and torque. Verification data have point deviations for most output parameters. It is clear from the graphs that the GFFS has a larger point deviation than the LHS. For both TSS, the model had the least accurate predictions for torque ripple.

Another indicator of all models accuracy were two statistical coefficients for machine learning: Pearson's R [13] and R^2 [14]. Pearson's R coefficient determines the strength of the linear relationship between two data sets. R^2 is a statistical measure of fit that indicates how much of the variability in the dependent variable is explained by the independent variable(s) in the regression model. Predicted and simulated data were used as inputs for the coefficients. Ideally, both coefficients should be equal to 1 or as close to it as possible. For the training data, the coefficients were very close to 1, indicating very small errors. For the verification data, they are shown in Table II. LHS again has better overall results.

Further indicators of all model's accuracy are the error analysis between the prediction and simulation values for the training and verification data. The training data exhibits very small errors, therefore error analysis is presented only for the verification data using typical statistical coefficients, as shown in Table III. For GFF, the absolute difference between the minimum and maximum error is larger than for LHS. Although in some cases the maximum and minimum errors are larger for the LHS, the difference is more decisive. A crucial coefficient is a variance, which shows how large the extremes are between the errors. For GFF, its values are significantly higher and show larger extremes between errors. On average, the LHS shows better results with smaller prediction errors.

The last important surrogate models parameter is uncertainty. In general, for a high degree confidence model, the value of uncertainty should be as low as possible. Here it is not possible to visualize it as in Fig. 2. Therefore, surrogate model uncertainties interpolation was performed for all combinations of input parameters. This interpolation acts as a dimensional reduction (from 4D to 2D) and serves only as an indicator of the uncertainty distribution, not as factual information. However, it is possible to see at which location each parameter in the model is most certain or uncertain. Two main examples for the torque and power factor between active length and

TABLE II
STATISTICAL COEFFICIENTS FOR ACCURACY COMPARISON OF SURROGATE MODELS BETWEEN GFF AND LHS FOR TSS

Coefficient	TSS	Parameter			
		Torque	ElMag. efficiency	Power factor	Torque ripple
Pearson's R	GFFS	0.9999684	0.9981351	0.9980106	0.9767286
	LHS	0.9999733	0.9994358	0.9992861	0.9868065
R^2	GFFS	0.9999367	0.9961802	0.9959937	0.9421553
	LHS	0.9999450	0.9986025	0.9984511	0.9722887

TABLE III
STATISTICAL COEFFICIENTS FOR VERIFICATION DATA ERROR COMPARISON BETWEEN GFF AND LHS FOR TSS

Coefficient	TSS	Parameter			
		Torque	ElMag. efficiency	Power factor	Torque ripple
Maximum (%)	GFFS	1.41	3.227	5.546	20.094
	LHS	1.454	1.032	4.347	10.922
Minimum (%)	GFFS	-1.938	-11.241	-9.077	-19.121
	LHS	-0.444	-3.395	-2.370	-21.230
Mean (%)	GFFS	-0.008	-0.222	0.085	0.149
	LHS	0.063	-0.136	0.062	-0.861
Median (%)	GFFS	0.021	0.032	0.100	0.140
	LHS	-0.007	-0.013	-0.076	-0.367
Variance ($\%^2$)	GFFS	0.187	2.638	2.271	30.892
	LHS	0.111	0.538	0.861	19.280
Standard deviation (%)	GFFS	0.432	1.624	1.507	5.558
	LHS	0.333	0.733	0.928	4.391

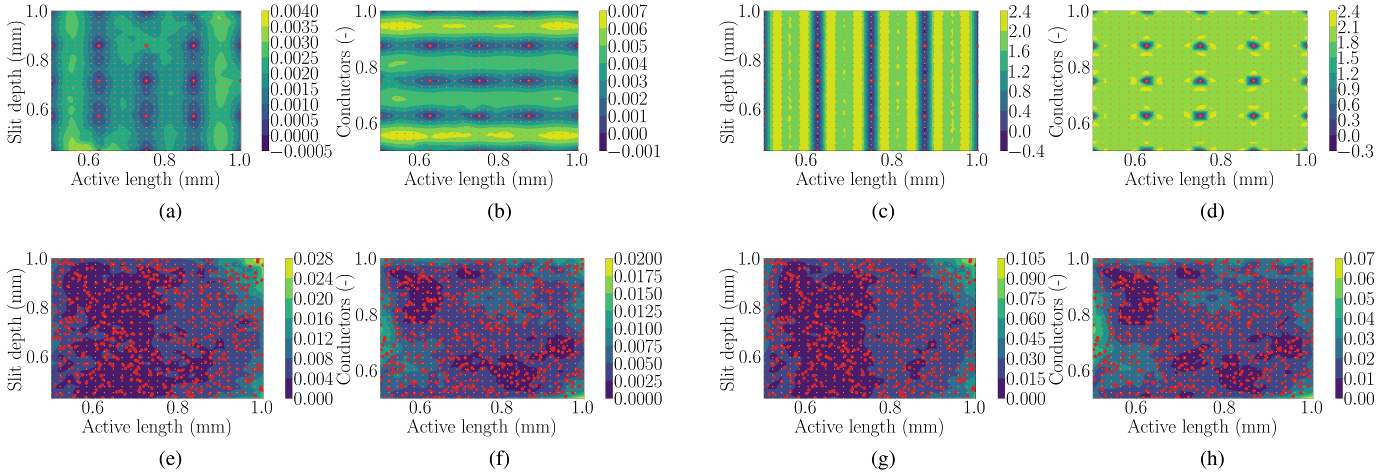


Fig. 7. Selected interpolated uncertainties of the constructed surrogate models for the machine's torque using GFFS (a), (b) and LHS (e), (f) TSS and machine's power factor using GFFS (c), (d) and LHS (g), (h) TSS.

slit depth/number of conductors will be presented here, see. Fig. 7. In the 2D plot, the color scale indicates the size of the uncertainty, while the red dots indicate the training data of the surrogate models. For torque, both TSS have relatively high prediction certainty, as can be seen in Fig. 6. GFFS is more uncertain in the model space, while LHS is more uncertain at the edges of the model space. The situation is quite different for the power factor, with the LHS model being relatively certain (except at the edges of the space), while the GFFS model is very uncertain between training points. Such high uncertainty for GFFS also applies to electromagnetic efficiency. Obviously, the LHS shows much better results here and overall compared to the GFFS. However, it should be pointed out that these are initial training data. In the case of active learning, the situation might look completely different. Nevertheless, the results show LHS generally exhibits better prediction and performance for surrogate models in this study.

IV. Conclusion

The machine learning technique for modeling of an IM was presented in this paper. A special case of machine learning technique, surrogate modeling, was adopted. For surrogate modeling, GPM regression using the Matérn kernel was employed, where 3 considered TSS were compared with each other. BBS was excluded due to the small sample size. Based on the results of this study, it was concluded that the best TSS was LHS, which showed much better results compared to the GFFS. Although the calculations were conducted on a solid rotor IM, the presented methodology can be applied to any type of electrical machine with a high potential for significant time-saving in the design of electrical machines.

References

[1] M. Yilmaz and P. T. Krein, "Capabilities of finite element analysis and magnetic equivalent circuits for electrical machine analysis and design," 2008 IEEE Power Electronics Specialists Conference, Rhodes, Greece, 2008, pp. 4027-4033, doi: 10.1109/PESC.2008.4592584.

[2] S. Silber, W. Koppelstätter, G. Weidenholzer, G. Segon and G. Bramerdorfer, "Reducing Development Time of Electric Machines with SyMSpace," 2018 8th International Electric Drives Production Conference (EDPC), Schweinfurt, Germany, 2018, pp. 1-5, doi: 10.1109/EDPC.2018.8658312.

[3] R. Alizadeh, J.K. Allen, and F. Mistree, "Managing computational complexity using surrogate models: a critical review," Res Eng Design 31, pp. 275–298, 2020, doi: 10.1007/s00163-020-00336-7

[4] K. McBride and K. Sundmacher, "Overview of Surrogate Modeling in Chemical Process Engineering," Chemie Ingenieur Technik, vol. 91, pp. 228-239. doi: 10.1002/cite.201800091.

[5] E. Schulz, M. Speekenbrink, and A. Krause, "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions," Journal of Mathematical Psychology, vol. 85, pp. 1-16, 2018, doi: 10.1016/j.jmp.2018.03.001.

[6] A. Denzel and J. Kästner, "Gaussian process regression for geometry optimization," The Journal of Chemical Physics, vol. 148, no. 9, pp. 094114, 2018, doi:10.1063/1.5017103.

[7] M. Buragohain and Ch. Mahntan, "A novel approach for ANFIS modelling based on full factorial design," Applied Soft Computing, vol. 8, no. 1, pp. 608-625, 2008, doi:10.1016/j.asoc.2007.03.010.

[8] A. Olsson, G. Sandberg, and O. Dahlblom, "On Latin hypercube sampling for structural reliability analysis," Structural Safety, vol. 25, no. 1, pp. 47-68, 2003, doi: 10.1016/S0167-4730(02)00039-5.

[9] G. Bramerdorfer, J. A. Tapia, J. J. Pyrhönen, and A. Cavagnino, "Modern Electrical Machine Design Optimization: Techniques, Trends, and Best Practices," in IEEE Transactions on Industrial Electronics, vol. 65, no. 10, pp. 7672-7684, Oct. 2018, doi: 10.1109/TIE.2018.2801805.

[10] A. Lee, "pyDOE: The experimental design package for python," pyDOE Design of Experiments for Python. https://pythonhosted.org/pyDOE/ (accessed January 31, 2023)

[11] J. Pyrhonen, J. Nerg, P. Kurronen and U. Lauber, "High-Speed High-Output Solid-Rotor Induction-Motor Technology for Gas Compression," in IEEE Transactions on Industrial Electronics, vol. 57, no. 1, pp. 272-280, Jan. 2010, doi: 10.1109/TIE.2009.2021595.

[12] "scikit-learn Machine Learning in Python," Gaussian Processes. https://scikit-learn.org/stable/modules/gaussian_process.html (accessed January 31, 2023)

[13] H. Akoglu, "User's guide to correlation coefficients," in Turkish Journal of Emergency Medicine, vol. 18, no. 3, pp. 91-93, 2018, doi: 10.1016/j.tjem.2018.08.001.

[14] L. J. Sanders, R. A. Russell, and D. P. Crabb, "The Coefficient of Determination: What Determines a Useful R 2 Statistic?," in Investigative Ophthalmology & Visual Science, vol. 53, no. 11, pp. 6830-6832, Oct. 2012, doi: 10.1167/iovs.12-10598.

Simulations of Loss Minimization Control Based on Motor Parameters for Indirect Field-oriented Control of Induction Motor

Tomáš Lažek
Department of power electrical
and electronic engineering, FEEC
Brno University of Technology
Brno, Czech Republic
ORCID: 0000-0003-1853-0416

***Abstract*—This paper discusses a simulation of a loss minimization algorithm for induction motor drive. This algorithm, which is based on motor parameters, varies the flux linkage to obtain the lowest possible losses. Verification of the algorithm's functionality was performed by simulations in MATLAB/SIMULINK environment. First, an induction machine model was built with the parameters of a real motor, taking into account magnetic saturation, Joule losses in both windings, iron losses, and mechanical losses. Then, simulations with indirect field-oriented control were performed. The simulations showed that there is an increase in efficiency in the low torque region.**

***Index Terms*—energy efficiency, indirect field-oriented control, induction motor drives, loss minimization**

I. Introduction

Nowadays, there is increasing pressure to reduce electricity consumption. Variable-speed drives with induction machines (IMs) are among the largest consumers of electricity. The biggest problem with IM drives is the low efficiency in the low torque region. There are several ways to increase efficiency.

The first method is called a search algorithm, which can be found in [1]. The principle is based on iteratively changing the flux to find the highest efficiency. The advantage is that the IM parameters are not required. On the other hand, the big disadvantage is the torque and speed ripple because the flux change is continuous.

The second method is based on knowledge of the IM parameters. In the literature this method is called loss minimization control [2]. The advantage is fast knowledge of the optimum flux linkage, but the big disadvantage is the need to know exactly the IM parameters, which can change during the operation.

The third category is called hybrid control and combines the advantages of both previous methods [3].

The goal of this paper is to demonstrate the procedure for determining a loss minimizing algorithm for indirect field-oriented control of an IM. The resulting algorithm will be simulated with real motor parameters. Finally, a comparison with the constant flux field-oriented control of IM will be made.

II. Induction machine model

A model of an IM with a squirrel cage is described in a synchronously (d - q) rotating frame. By changing the flux linkage constant, the indirect field-oriented control strategy can be varied to the stator flux orientation, the rotor flux orientation, or the air gap flux orientation. From the point of view of simplicity, the rotor flux oriented model is the most advantageous. The model is based on a commonly used model in the form of an inverse Gamma (Ꞁ) network with a resistor representing iron losses. A detailed development of the model with iron loss resistor can be found in [4]. The stator current in the d-axis i_{sd} is responsible for generating the magnetic field, while the stator current in the q-axis i_{sq} generates the torque. IM values were determined by the no-load test and the load test.

The IM model in form Γ-network is shown in Fig. 1 separately for the d-axis (a) and q-axis (b). The rotor flux is aligned in the direction of the d-axis, therefore $\Psi_{Rd} = \Psi_R$ and $\Psi_{Rq} = 0$. An iron loss resistance R_{Fe} is connected parallel to the stator branch. In Fig. 1 these elements are present 1) L_R the magnetization inductance on the rotor side, which is defined by $L_R = L_s k^2$, where L_s is the magnetization inductance obtained from the no load test and k is the coupling factor, 2) R_R the rotor resistance for the Ꞁ-network, which is defined as $R_R = R_r k^4$, where R_r is the rotor resistance obtained from the load test, 3) $L_{\sigma S}$ is the leakage inductance on the stator side expressed by $L_{\sigma S} = L_s(1-k^2)$. In steady state, the coils can be considered as a short circuit. Then, the stator voltage on the d-axis

$$u_{sd} = R_s i_{sd} + \omega_s L_{\sigma S} i_{sq}, \tag{1}$$

where ω_r is the rotor frequency (slip frequency), which is defined

$$\omega_r = \omega_s - p\omega_m. \tag{2}$$

ω_s is stator frequency, p is number of pole-pairs and ω_m is the mechanical speed on the motor shaft.

Then for the stator voltage in the q-axis

$$u_{sq} = R_s i_{sq} + \omega_s L_{\sigma S} i_{sd,\text{Ꞁ}} + \Psi_R \omega_s. \tag{3}$$

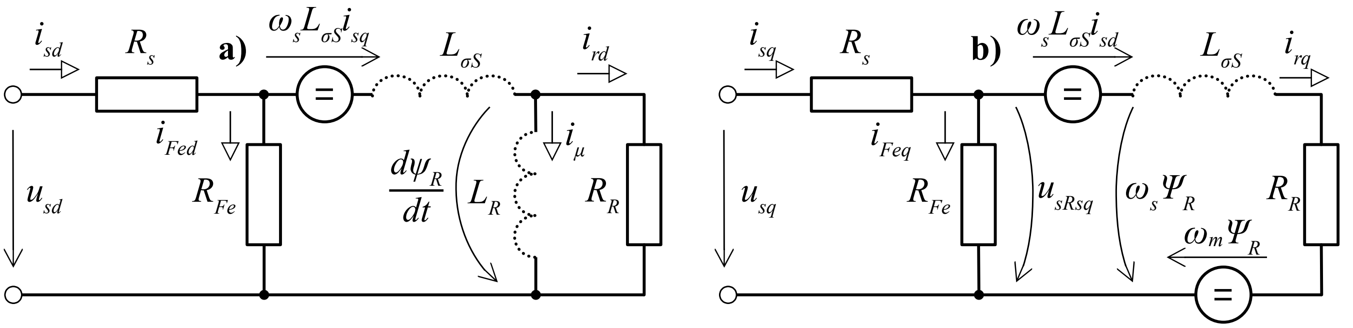


Fig. 1. IM equivalent circuit in form of Γ-network in: a) d-axis and b) q-axis.

For the calculation of the electromagnetic torque can be assumed that only the rotor current in the q-axis is torque-generating, then

$$T_e = \frac{3}{2} p \Psi_R i_{rq}. \tag{4}$$

The magnetization current can be expressed as

$$i_\mu = \frac{\Psi_R}{L_R}. \tag{5}$$

From Fig. 1 it can be seen that the rotor current in the d-axis is $i_{rd} = 0$. Then, for the stator current in the d-axis with aid (5) can be written

$$i_{sd} = i_{Fed} + i_\mu = i_{Fed} + \frac{\Psi_R}{L_R}. \tag{6}$$

From (4) can be expressed the rotor current in the q-axis, then for the stator current in the q-axis

$$i_{sq} = i_{Feq} + i_{rq} = i_{Feq} + \frac{2}{3} \frac{T_e}{p \Psi_R}. \tag{7}$$

III. Loss model

The loss model considers only the dominant losses, which are the Joule losses in the stator and rotor and the losses in the stator iron. The losses in the rotor iron are proportional to the rotor frequency. However, the rotor frequency is very small and therefore these losses are neglected. In addition, friction losses, windage, stray losses and power losses in the inverter are omitted as they have only a small impact on the total drive losses. Since the model is in a synchronous rotating frame, the power losses can be determined very simply.

Stator iron losses consist of eddy current losses and hysteresis losses. As shown in [4], the iron loss resistance may not be considered as constant, but can be modeled as a simple linear relationship between iron loss resistance and the stator frequency. Then the iron loss resistance is given

$$R_{Fe} = R_{Fe0} \frac{\omega_s}{\omega_0}, \tag{8}$$

where R_{Fe0} is the measured iron loss resistance at frequency ω_0. Only these two values must be experimentally determined.

In further calculations, it is necessary to express all quantities in terms of torque and mechanical speed. Since the rotor frequency is very small, it can be neglected in the expression for the stator frequency for simplicity. Then

$$\omega_s = p \omega_m. \tag{9}$$

Iron losses can be expressed in general

$$P_{Fe} = \frac{3}{2} \frac{u_{sRs}^2}{R_{Fe}} = \frac{3}{2} \frac{u_{sRs}^2 \omega_0}{R_{Fe0} \omega_s}. \tag{10}$$

To determine the voltage across the resistor R_{Fe}, Fig. 1 can be used. The voltage in the q-axis is defined $\Psi_R \omega_s + \omega_s L_{\sigma S} i_{sd}$. However, the voltage in the d-axis is considered negligible. Then the losses in iron are given

$$P_{Fe} = \frac{3}{2} \frac{p \omega_m \omega_0 \Psi_R^2 (1 + L_{\sigma S}/L_R)^2}{R_{Fe0}}. \tag{11}$$

The stator resistor and the rotor resistor flow currents in both axes, then in general can be written for the Joule losses

$$P_{js} = \frac{3}{2} R_s (i_{sd}^2 + i_{sq}^2), \tag{12}$$

$$P_{jr} = \frac{3}{2} R_r (i_{rd}^2 + i_{rq}^2). \tag{13}$$

The iron loss resistance is many times greater than the rotor resistance. Hence, the currents in the iron resistance can be neglected

$$i_{Fed} \approx i_{Feq} \approx 0. \tag{14}$$

Then the stator current in the q-axis can be expressed as a function of the torque

$$i_{sq} = i_{rq} = \frac{2}{3} \frac{T_e}{p \Psi_R}. \tag{15}$$

As mentioned above, if the rotor current in the d-axis is zero, then the stator current in the d-axis is formed only by the magnetizing current i_μ. With aid (15) the Joule losses in both windings can be expressed

$$P_{js} = \frac{3}{2} \frac{R_s \Psi_R^2}{L_R^2} + \frac{2}{3} \frac{R_s T_e^2}{p^2 \Psi_R^2}, \tag{16}$$

$$P_{jr} = \frac{2}{3} \frac{R_R T_e^2}{p^2 \Psi_R^2}. \tag{17}$$

IV. Loss minimization

The principle of loss minimization is based on optimizing the variables involved in torque generation. It follows from (4) that the rotor flux Ψ_R and the current i_{rq} generate torque. If the rotor flux linkage Ψ_R is decreased (increased) and the torque-generating current i_{rq} is increased (decreased), the same torque is maintained. The figure 2 shows the dependence of the power loss on the rotor flux linkage Ψ_R for different values of the load torque. At nominal torque, it can be seen that power losses cannot be reduced by changing the rotor flux linkage. However, as the load decreases, the power losses can be reduced by decreasing the rotor flux.

Determining the analytical expression for the optimal steady-state flux linkage is very simple and can be found by solving

$$\frac{\partial}{\partial \Psi}\{P_{Fe} + P_{js} + P_{jr}\} = 0. \quad (18)$$

After substituting (15), (16) and (17), an analytical expression for the optimal rotor flux can be obtained

$$\Psi_{R,opt} = \sqrt{T_e}\sqrt{\frac{2L_R}{3p}}\sqrt[4]{\frac{R_{Fe0}(R_s + R_R)}{\omega_m \omega_0 p L_R^2 A^2 + R_{Fe0} R_s}}, \quad (19)$$

where $A = 1 + L_{\sigma S}/L_R$.

V. Simulations

Simulations were performed in the MATLAB/SIMULINK environment. The parameters of a real IM were used, on which a load test and a no load test were performed. Since identification is not the focus of this paper, all the parameters are given in the Appendix. Furthermore, for the simulation, the inverter supply is not considered.

A. Induction machine model

The IM model is developed in a stationary (α - β) frame in the form of the Γ-network. For the general model, which can be found in [5], a resistance is added to the stator branch representing the losses in iron. A block for calculating mechanical losses is also included in the model. Furthermore, the magnetization inductance is not modeled as a constant but as a dependence on the flux linkage.

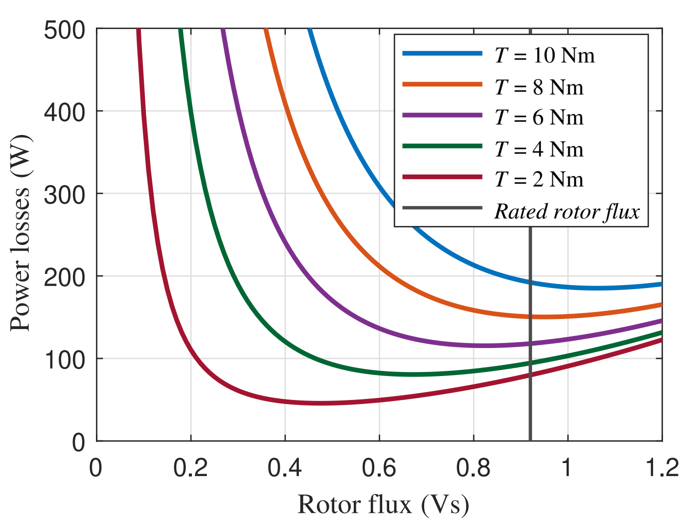


Fig. 2. Dependence of power losses on rotor flux.

Iron losses are modeled separately for hysteresis and eddy current losses. These losses can be expressed as functions of the linkage flux and the stator frequency

$$P_{Fe} = k_h \frac{\omega_s}{2\pi}\Psi_s^{n_h} + k_e\left(\frac{\omega_s}{2\pi}\right)^2 \Psi_s^2, \quad (20)$$

where k_h is the hysteresis loss coefficient, n_h is the hysteresis loss exponent and k_e is the eddy current loss coefficient. Both of these coefficients can be obtained by experimental measurement. To calculate the current i_{Fe}, the conductivity in iron G_{Fe} is expressed from (20). Then i_{Fe} can be obtained

$$i_{Fe} = U_{s,Rs} G_{Fe}, \quad (21)$$

where $U_{s,Rs}$ is the voltage across the iron loss resistor.

Mechanical losses are divided into two components, friction losses, and windage losses. Friction losses have a linear dependence on mechanical speed, and windage losses, being of aerodynamic origin, are dependent with the third power on mechanical speed. The two losses can be simply added together to yield

$$T_{mech} = k_f \omega_m + k_w \omega_m^3, \quad (22)$$

where k_f is the friction loss coefficient and k_w is the coefficient of windage losses. The mechanical loss torque T_{mech} are then subtracted from the internal electromagnetic torque T_e of the motor.

Since the magnetizing inductance depends on the linkage flux, it is not considered as a constant in the model. Instead, a fifth-degree polynomial is introduced, whose coefficients can be obtained from the no load test. Then

$$L_R = L_s k^2 = \sum_{i=0}^{5} a_i \Psi_s^i k^2, \quad (23)$$

where a_i are polynomial coefficients, which are given in the appendix.

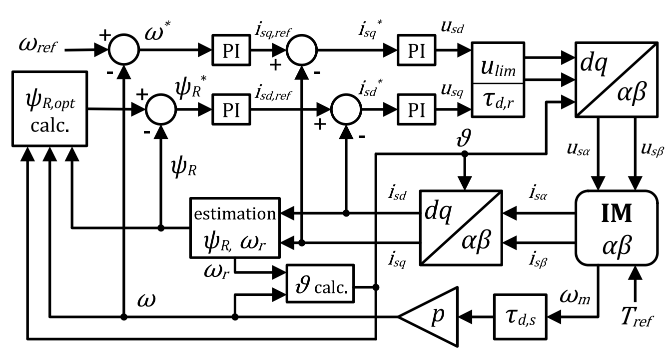


Fig. 3. Simplified IM field-orientend control scheme with optimal rotor flux calculation for Γ-network.

B. Indirect field-oriented control

The classic indirect field-oriented control with orientation on rotor flux, which can be found in [4] is chosen. As mentioned above, this is the simplest way for decoupling the d-axis and q-axis components. Fig. 3 shows a simplified control scheme. Compared to the classical indirect field-oriented control method, the reference rotor flux is not constant, but is calculated using (19). Furthermore, some characteristics of the real inverter are simulated, such as response time ($\tau_{d,r}$) and voltage limitation ($u_{lim} = \sqrt{2} \cdot 230$ V), which corresponds to the maximum achievable voltage of the power grid. The speed sensor delay ($\tau_{d,s}$) is also included in the simulation.

C. Simulations results

In the Fig. 3, the dependence of a) efficiency, b) rotor flux, c) input power and d) iron losses on torque can be seen. The solid lines correspond to the control with optimal rotor flux, and the dashed lines correspond to the control with constant rotor flux. The parameter is three different mechanical speeds ($\omega_m = \omega_n/3$, $\omega_m = 2\omega_n/3$, $\omega_m = \omega_n$). In the first graph in Fig. 3 a) the comparison of the efficiencies are sketched. It can be seen that the largest energy savings occur at low torque region. However, no energy savings occur in the nominal torque region. The second graph in Fig. 3 b) shows the nominal rotor flux compared to the optimal rotor flux. There is a significant reduction in the low torque region. In the nominal torque region, the rotor flux exceeds the nominal value. The third graph in Fig. 3 c) shows the input power. In the low torque regions, a reduction in input power can be seen for the control with optimal rotor flux. The last graph in Fig. 3 d) shows the comparison of iron loss. Again, it can be seen that there is a reduction of the iron losses in the low torque region for the control with optimal rotor flux. This is the main reason energy savings occur.

VI. Conclusion

In this paper, an algorithm for loss minimization in IM drive has been presented and simulated. The algorithm was implemented in an indirect field-oriented control (synchronous d-q frame), while the power supply from a real inverter was not considered. The IM model is developed in a stationary (α - β) frame with the parameters of the real IM obtained from measurements.

The simulations show that the greatest improvement in efficiency occurs in the low torque region. The efficiency is

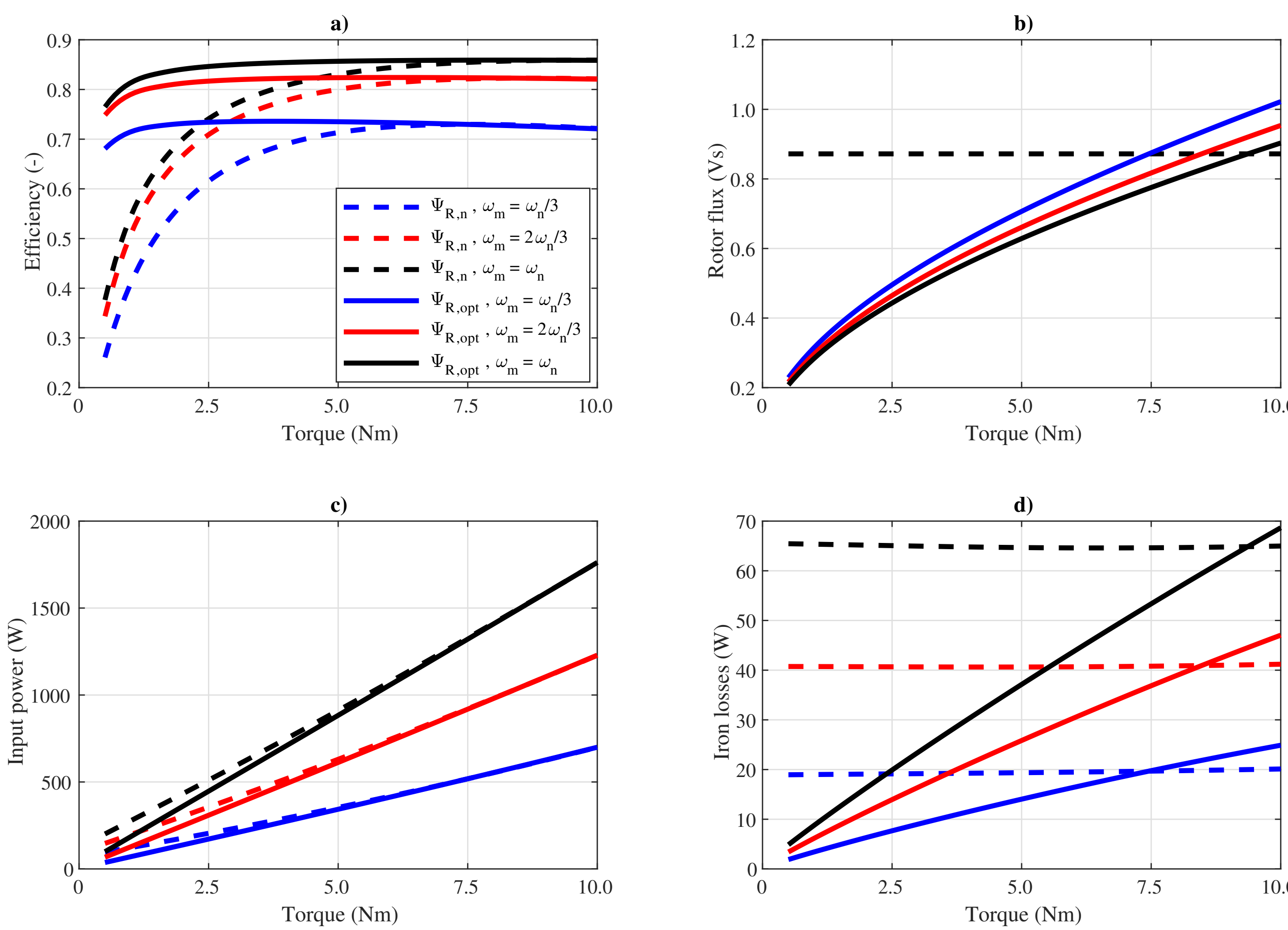


Fig. 4. Dependence of efficiency, rotor flux, power losses, and iron losses on torque at different speeds. The legend is the same for all four plots.

almost twice as high as in conventional constant flux field-oriented control. Furthermore, in this region, the rotor flux linkage is reduced, and thus the iron losses are reduced. On the other hand, in the nominal torque region, the efficiencies are similar.

Simulations have shown that the algorithm can reduce losses and thus increase efficiency. Therefore, the algorithm can be implemented in a real drive with IM, which will be a continuation of this research.

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 Ivo Pazdera as my research supervisor for his valuable comments and help while writing this paper.

APPENDIX

The IM nameplate data are as follows: $P_n = 1500$ W; $U_n = 400$ V; $f_n = 50$ Hz; $I_n = 3.15$ A; $n_n = 1445$ rpm; $p = 2$; $T_n = 9.9$ N·m; $\eta = 85.3$ %. Additional parameters were identified experimentally with the following results: $R_s = 3.082\ \Omega$; $L_{\sigma R} = 0.0227$ H; $R_r = 2.78$ Ω; $J = 5.9276$ g·m^2;$k = 0.96$; $k_h = 1.164$ W·s·T$^{-n_h}$; $n_h = 1.84$; $k_e = 6.23 \cdot 10^{-3}$ W·s·T^{-2}; $k_f = 4.85 \cdot 10^{-2}$ W·s·rad^{-1}; $k_w = 1.496 \cdot 10^{-7}$ W·s^3·rad^{-3}; $a_0 = 0.3115$ H; $a_1 = 0.3529$ H·V^{-1}·s^{-1}; $a_2 = -0.9841$ H·V^{-2}·s^{-2}; $a_3 = 1.06$ H·V^{-3}·s^{-3}; $a_4 = -0.5249$ H·V^{-4}·s^{-4}; $a_5 = 0.0728$ H·V^{-5}·s^{-5}.

REFERENCES

[1] H. Bizhani, S. M. Muyeen, F. R. Tatari, F. Gao and H. Geng, ”Comparative Analysis of Search Algorithm based Loss Minimization Techniques used in Vector Controlled Induction Motors,” *2020 2nd International Conference on Smart Power And Internet Energy Systems (SPIES), 2020*, pp. 386-390, doi: 10.1109/SPIES48661.2020.9243036.

[2] M. N. Uddin and S. W. Nam, ”New Online Loss-Minimization-Based Control of an Induction Motor Drive,” in *IEEE Transactions on Power Electronics*, vol. 23, no. 2, pp. 926-933, March 2008, doi: 10.1109/TPEL.2007.915029.

[3] C. Chakraborty and Y. Hori, ”Fast efficiency optimization techniques for the indirect vector-controlled induction motor drives,” in *IEEE Transactions on Industry Applications*, vol. 39, no. 4, pp. 1070-1076, July-Aug. 2003, doi: 10.1109/TIA.2003.814550.

[4] N.P. Quang and J.A. Dittrich. *Vector Control of Three- Phase AC Machines*. Berlin, Germany: Springer-Verlag, 2008.

[5] A. Veltman, D. Pulle, R. De Doncker *Fundamentals of Electrical Drives*, Springer Netherlands, 2007

vitesco TECHNOLOGIES

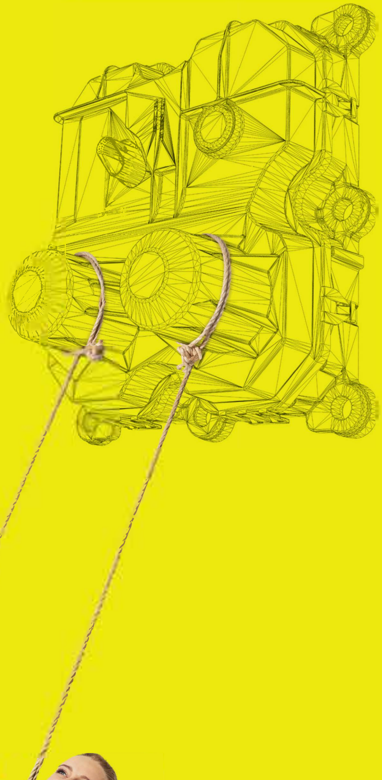
PŘEMÝŠLÍŠ, CO DÁL PO ŠKOLE?

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!

Advancements in the development of a proportional counter measuring system

Ondřej Kolář
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
ORCID:0000-0003-4187-4890

Michal Kubíček
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
ORCID:0000-0001-8247-8013

***Abstract*—This paper is an interim report of theory and continuous work on the proportional counter measuring system. The first part of the article covers general principles and properties of several types of electronic detectors for measuring the ionizing radiation fields including the proportional counters, to highlight their positives and negatives. The following part discusses the processing of proportional counter output signal. Next, the conducted measurement of a neutron source is described, and its results are shown. Based on the gained experience, the simulation and design of a new preamplifier for the measuring system was carried out.**

***Index Terms*—ionizing radiation, proportional counters, measurement system, low noise amplifier, transimpedance amplifier**

I. Introduction

Ionizing radiation is utilized in many fields of modern science and research, but it is also used in power generation or commercially. It is a set of physical phenomena which can, by definition, ionize a material or medium that has been exposed to it. Depending on many factors, it can have potentially harmful effects on life organisms or inanimate objects. Therefore, the need for trustworthy and precise characterization of ionizing radiation fields is crucial.

The ionizing radiation can be described as a stream of moving particles. The number of particles traveling through a given area per unit of time is called particle flux, and it indicates the intensity of a given radiation field. The next important parameter is the energy carried by every individual particle since it indicates the amount of ionization that can be caused by the said particle. Last but not least, the type of particle is also very important. Based on all of these parameters combined, we can estimate the radiation behavior, its degree of danger, and also its source can be identified or studied.

II. Detectors

Various detector types are used for the measurement of ionizing radiation. The choice of type depends on many factors and will vary for different applications. This article covers only some of the electronic detectors, but other types exist as well [1].

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.

The most important aspect of every detector is its measurement capabilities. They differ in spectral and particle-type sensitivity and resolution, the maximum detectable intensity, or the output signal strength [2].

A. Scintillation counters

Scintillators are a group of materials which emit visible light when irradiated with ionizing particles. Every incident particle transfers its energy to the material which then re-emits it in the form of a visible light impulse. These impulses are very short and dim, so for their conversion to an electric signal, photomultiplier tubes (PMTs) are often used [2]. The more energy the incident particle dissipates, the more photons of visible light are generated. This results in the output electric impulse being proportional to the energy of the originating particle. Thanks to this, scintillation counters are widely used for ionizing radiation spectrometry [2].

The scintillator could be of various materials [3]. Solid scintillators often come in the form of cylindrical crystals with a protective cover. Liquid scintillators must be contained in specialized sealed containers. Both of them must be made in such a way as to block all of the light from surrounding space, as any of it would overwhelm the scintillation photons and could potentially damage the PMT if turned on. They must also have an optical window for coupling with the PMT. This optical coupling has to be well-made to be as transparent as possible. The PMT serves two purposes. It transforms the optical signal into an electric signal, which then gets amplified for further processing in subsequent electronics [1] [2].

B. Gaseous ionization detectors

Gaseous ionization detectors are a group of ionizing radiation detectors. All of them share the same general mechanical design, which is depicted in Fig. 1. Their electrically conductive shell is connected to the negative terminal of a voltage source. The positive terminal is connected to the anode, which usually consists of a thin wire in the center of the chamber. The inside of the chamber is filled with a gas or a gas mixture [2].

When an ionizing particle enters the detector, it ionizes the gas fill along its trajectory. The more energy the incident particle dissipates, the more of the gas gets ionized. These ions

DOI 10.13164/eeict.2023.198

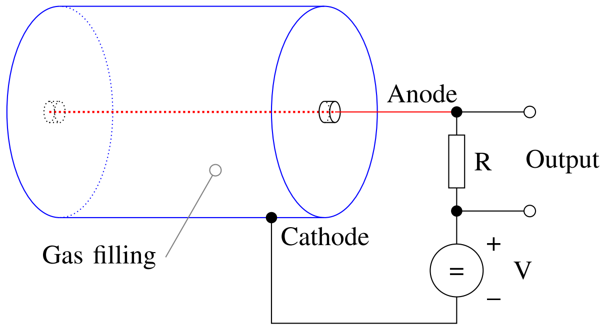


Fig. 1. Diagram of a simple cylindrical gaseous ionization detector [2]

start to move toward the corresponding electrodes as they are accelerated by the electric field. The acceleration is directly proportional to the intensity of the electric field, which is determined by the shape of both electrodes and the voltage supplied to them. When any ions finally reach the electrodes, they manifest as a current impulse at the output of the detector [2].

The optimal physical shape of the detector is chosen with respect to the intended region of operation, which then determines properties of the output signal. For a given detector shape, the region of operation depends on the supplied voltage as seen in Fig. 2. The chart shows the relation between the supply voltage and the output signal strength for two incident particles of different energies.

1) Geiger-Müller tubes: are a well-known group of gaseous ionization detectors designed to work in the region of operation of the same name. They are usually of a thinner cylindrical shape. Due to this, a high electric field intensity can be achieved without the need of a very high supply voltage. This strong electric field accelerates the primary free ions, created by an incident ionizing particle, so much that they hit other gas particles. This causes secondary ionization and causes a chain reaction that forms multiple Townsend avalanches

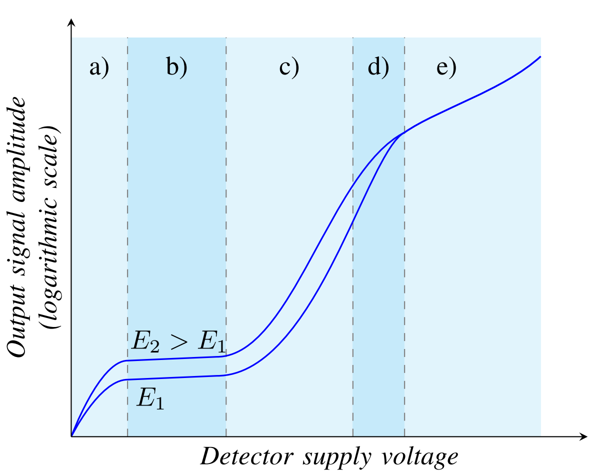


Fig. 2. Regions of operation of gaseous ionization detectors (a: ion recombination; b: ion saturation; c: proportional; d: limited proportionality; e: Geiger-Müller) [2]

throughout most of the detector interior. This means that the amplitude of the output current pulse is relatively high, therefore the subsequent processing of the signal is easy and does not require very sophisticated electronics. However, since the detector gas fill becomes saturated with the avalanches by any incident particle regardless of its properties, only the intensity of ionizing radiation fields can be measured with the Geiger-Müller tubes by counting the number of pulses per unit of time [2].

2) Ionization chambers: are a group of gaseous ionization detectors that work in the *ion saturation region*. On the contrary of the Geiger-Müller tubes, in this region, the primary ions are accelerated sufficiently so that neither their recombination is possible nor they could cause secondary ionization. This means that their amount is the same as the number of ions reaching the detector electrodes. Therefore, the resulting output current pulse is directly proportional to the energy dissipated by the incident ionizing particle. The downside is the small amplitude of such signal, so the requirements for a very high amplification and a low noise properties of the subsequent processing electronics are inevitable [2].

3) Proportional counters: are a group of detectors working in the *proportional region*, which combines the benefits of both previously mentioned regions. As the plot in Fig. 2 shows, the output signal amplitude remains proportional to the energy of an incident particle. Additionally, the amplitude also increases with higher detector supply voltages. This is caused by the electric field intensity near the anode in the detector center, which is higher than the minimum intensity needed for Townsend avalanche formation. The avalanche region is kept only within a small radius around the anode, as seen in Fig. 3, to ensure the formation of small nonoverlapping avalanches from primary electrons. This process multiplicates the number of secondary ion pairs and is called *gas amplification* [2]. As the final number of ions is much larger, the output signal amplitude is higher, and so is the signal-to-noise ratio. Therefore, the low-noise requirements of processing electronics are less demanding than those of ionization chambers.

III. Proportional counter signal processing

An output signal of a proportional counter comes directly in the form of an electric current pulse. Although its amplitude is higher, relative to an ionization chamber, it is still too weak for any direct processing. Therefore, it has to be amplified first. The specialized proportional counter preamplifiers often serve multiple purposes. In addition to amplification, they inject the supply voltage into the detector. They also transform the current signal to a voltage signal.

A. Particle energy measurement

Commercial preamplifiers for ionizing radiation spectrometry often come as charge sensitive amplifiers, which are basically integrating transimpedance amplifiers. Therefore, their amplification factor (gain) is given in units of $\mathrm{mV/pC}$ [4]. The integration of the signal comes with two benefits. The first is noise filtering, since the integrator behaves like

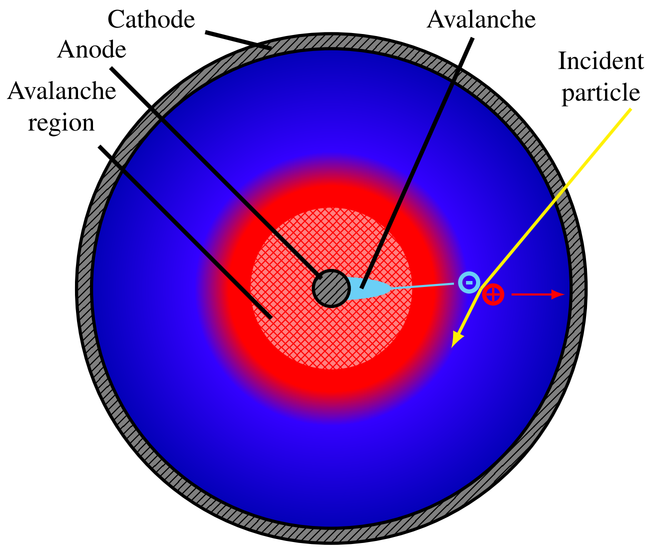


Fig. 3. Example of a single ion pair interaction in a cylindrical proportional counter [2] [5]

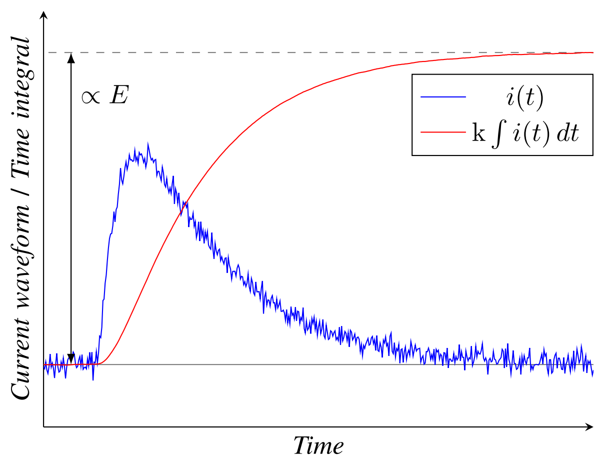


Fig. 4. Example of an artificial noisy current pulse and its numerical integration

a low-pass filter. The second benefit comes with the fact that the current $i(t)$ through the detector terminals in any given moment is proportional to the rate of ions reaching the electrodes. By integrating this current pulse, it results in the total amount of charge generated by the incident particle which is proportional to the energy E dissipated by it. This simplifies further processing of the signal because the particle energy, which is of interest when using the proportional counter for spectroscopy, can be determined by measuring the height of the voltage step. To show both of these properties of charge sensitive preamplifiers, an example noisy pulse was generated and its integral (scaled by an appropriate constant k) computed using MATLAB and plotted to Fig. 4.

B. Pulse shape discrimination

When measuring ionizing radiation field, it can be composed of multiple types of particles simultaneously. One example of

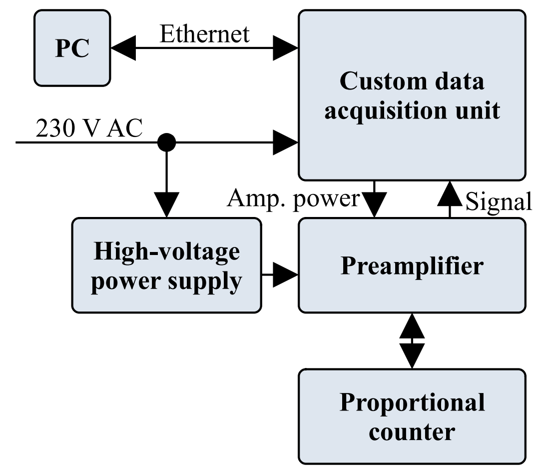


Fig. 5. Diagram of the measuring system

these mixed fields can be neutron radiation, which is often accompanied by gamma radiation, as the nuclear reactions usually produce both of them. For a proper characterization of such a field, the ability to separate them is important.

Pulse shape discrimination are various methods of signal analysis for classification of an incident particle type. For the scintillation counters, described in II-A, several of these methods are established and commonly used [6] [7]. However, the usual combination of charge sensitive preamplifiers and shaping amplifiers used with the proportional counters are deteriorating the original pulse shape. This makes them unusable for subsequent pulse shape analysis. The use of a non-integrating transimpedance amplifier is more suitable for the application of a pulse shape discrimination method [8].

IV. Measurement

The custom developed measuring system used for data acquisition from proportional counters consists of several parts shown in Fig. 5. Its detailed description is covered in [9].

Using the system, the first real measurement of a radiation field was conducted at the Department of Electrical Power Engineering, FEEC, BUT. As a source of radiation the Americium-beryllium neutron source AS010/15, manufactured by Eckert & Ziegler Cesio s.r.o. was used. The commercial charge sensitive preamplifier used was the CSP10 manufactured by FAST ComTec GmbH. Its gain is $1.4\,\mathrm{V{\cdot}pC^{-1}}$. The spherical proportional counter used for the measurement was manufactured by LND, Inc. with the serial number 270133 and it was supplied with 2500 V.

The detector was placed in the shielded room on top of a fixture in which the neutron source is securely kept. However, the number of detected events was too small to acquire a reasonable amount of data. To increase the number of detected particles, the neutron source was taken out of the fixture and placed at a distance of 20 cm from the detector. This drastically increased the number of detected events. A photo of the detector and source placement is in Fig. 6.

Fig. 6. Placement of the detector (left) and the neutron source (right) during the measurement

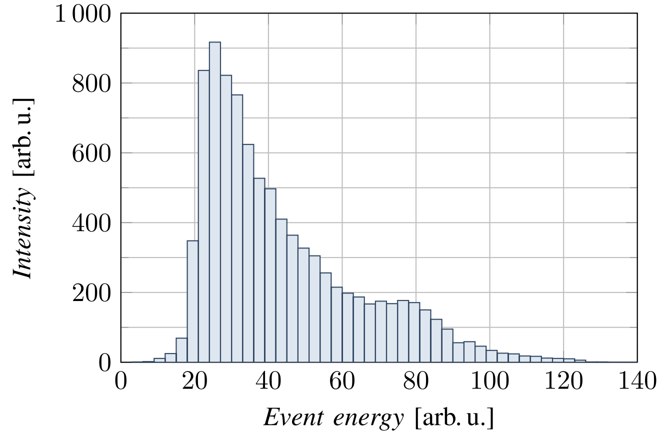


Fig. 8. The measured energy spectrum

The measurement was run for 10 minutes and 5 seconds, during which the number of captured events was 15 019. The average detection rate was approximately 24.8 events per second. The saved binary data was then loaded into MATLAB for processing.

Since the system saves the signal in raw form, a simple moving average filter was applied to filter out the noise first. The difference between the raw and the filtered signal can be seen in Fig. 7. The use of the integrating preamplifier facilitates easy computation of a measured energy spectrum by calculating the height of every impulse as described in Section III-A and plotting the values to a histogram (Fig. 8).

V. Preamplifier design

Following the measurement, it was decided that a custom non-integrating amplifier for the measuring system will be designed and built. The reason for this is that the manufacturer prevents any modifications to the commercial CSP10, as the printed circuit board of the amplifier module is obfuscated by encasement in a black epoxy resin.

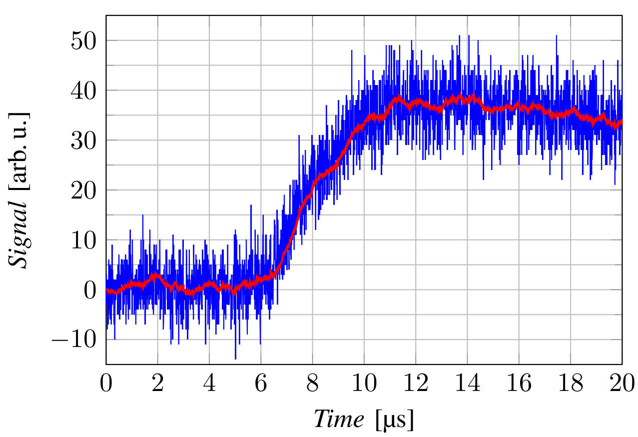


Fig. 7. Raw waveform of a captured event (blue) and its filtered form (red)

As a core for the preamplifier, the AD8099 integrated circuit was chosen. It features an ultra-low noise operational amplifier with high-speed capabilities given by its gain-bandwidth product of 3.8 GHz [10]. The manufacturer Analog Devices, Inc. provides useful designer tools [11], including the Photodiode Circuit Design Wizard and SPICE model of the AD8099.

A. First stage design

The Photodiode Circuit Design Wizard is an online tool for a simple and quick design of a photodiode transimpedance amplifier. The photodiode is simulated as a square waveform current source with its capacitance and shunt resistance in parallel. The values of all parameters can be set according to the required design. To model a proportional counter, the peak current was set at 100 µA and the capacitance to 50 pF. The shunt resistance was set to 100 MΩ as this is a value of a resistor cascade through which the proportional counter is supplied with high voltage.

In the second tab of the tool, the AD8099 was selected. Various values of the required output signal parameters were then tried, and their influence on the estimated circuit characteristics was observed. The reasonable balance was found by setting the required bandwidth of 100 MHz, peak output voltage 100 mV, and peaking capacitor value 2.4 pF. The tool then estimated the output signal-to-noise ratio of 37.1 dB.

B. Simulation

With the SPICE model of the AD8099 available, the circuit was transferred to the LTspice simulator. For better utilization of the maximum ± 1 V range of the ADC in the acquisition unit, a second inverting amplifier stage with the voltage gain of 10 was added to the design. To ensure the modularity of the measuring system, both stages will be constructed separately, each of them in its own casing. This requires proper impedance matching to 50 Ω coaxial cables used for interconnection. The final simulation schematic is shown in Fig. 9. Running the noise analysis of the circuit in the range of 1 Hz to 1 GHz revealed its frequency characteristics, which are plotted in Fig. 10. The resulting bandwidth ranges approximately from

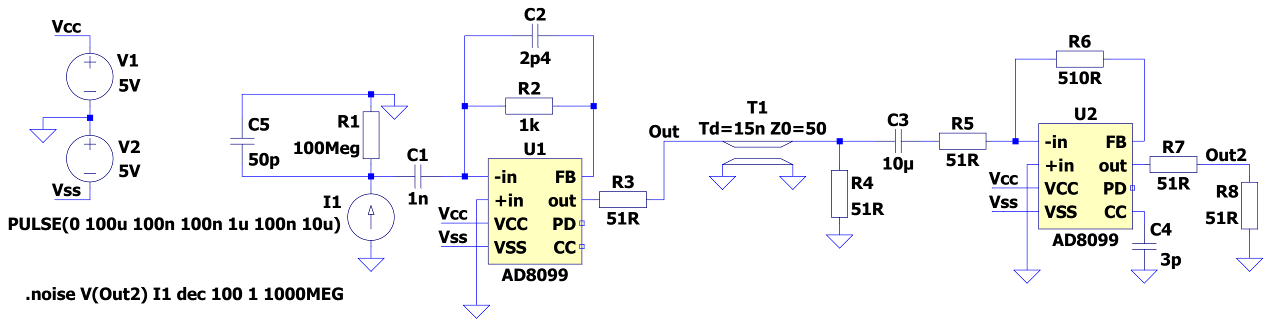


Fig. 9. The designed preamplifier simulation schematic

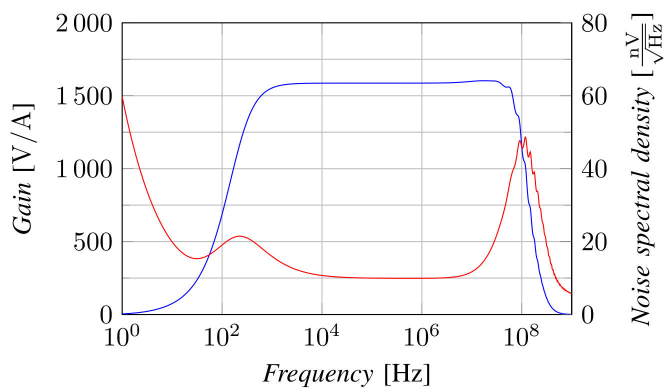


Fig. 10. The simulated frequency characteristics - gain (blue), noise spectral density (red)

246 Hz to 96.8 MHz with a gain of 1.59 V/mA. The total output RMS noise is 653.4 µV.

C. Physical design

The physical design of the preamplifier followed. To lower the cost of manufacturing, only a single universal version of a 4-layer printed circuit board (PCB) design was created using KiCad software. The board contains only one amplifier stage with all of the passive component footprints needed for populating either of the stages.

The power for the AD8099 amplifier is provided by the measuring unit which outputs a symmetric voltage of ± 15 V from its internal switching power supply. Since the voltage is quite noisy and too high for the amplifier, it is filtered by an LC filter and lowered to the ± 5 V by a dual symmetric low noise LDO regulator LT3032-5.

To ensure the low noise properties, the PCB was designed to fit inside an extruded aluminum enclosure, and on top of that, the AD8099 circuitry is placed under an additional SMT EMC shield to ensure suppression of any noise which could be brought inside of the amplifier enclosure with the power supply voltage.

VI. Conclusion

Taking into account the theory and experience gained from the conducted measurement, a custom modular preamplifier for the measuring system was designed. The design was verified using the LTspice simulator, which showed promising results. However, the simulation did not account for all real-world inaccuracies, such as parasitic parameters of the components used or especially the parasitic capacitance of the PCB, therefore the results might not be accurate and will need to be verified.

So, as a next step, a physical prototype of the preamplifier is about to be built, tested, and fine-tuned if necessary. At the time of writing this paper, the prototype is already in the process of production. The experimental results are therefore expected to be known soon.

References

[1] V. Ullmann, "Detekce a spektrometrie ionizujícího záření", AstroNukl-Fyzika.cz.

[2] G. F. Knoll, "Radiation Detection and Measurement," 3rd ed. Michigan: John Wiley & Sons, 2000.

[3] B. D. Milbrath, A. J. Peurrung, M. Bliss, and W. J. Weber, "Radiation detector materials: An overview", *Journal of Materials Research*, vol. 23, no. 10, pp. 2561-2581, 2008.

[4] FAST ComTec. (2023). *CSP10 Charge Sensitive Preamplifier* [Online]. Available: https://www.fastcomtec.com/amp/csp10-charge-sens-preamp

[5] T. Neep, I. Katsioulas, P. Knights, J. Matthews, K. Nikolopoulos, and R. J. Ward, "A simulation framework for Spherical Proportional Counters", in *Proceedings of 40th International Conference on High Energy physics — PoS(ICHEP2020)*, 2021, p. 917.

[6] G. Liu, M. D. Aspinall, X. Ma, and M. J. Joyce, "An investigation of the digital discrimination of neutrons and gamma rays with organic scintillation detectors using an artificial neural network", *Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip.*, vol. 607, no. 3, pp. 620-628, 2009.

[7] I. A. Pawełczak, S. A. Ouedraogo, A. M. Glenn, R. E. Wurtz, and L. F. Nakae, "Studies of neutron– pulse shape discrimination in EJ-309 liquid scintillator using charge integration method", *Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip.*, vol. 711, pp. 21-26, 2013.

[8] T. J. Langford, C. D. Bass, E. J. Beise, H. Breuer, D. K. Erwin, C. R. Heimbach, and J. S. Nico, "Event identification in 3He proportional counters using risetime discrimination", *Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip.*, vol. 717, pp. 51-57, 2013.

[9] O. Kolář, "Systém pro měření s proporcionálními detektory", Diplomová práce, Brno, 2022.

[10] Analog Devices, "Ultralow Distortion, High Speed, 0.95 nV/Hz Voltage Noise Op Amp", AD8099 datasheet, 2016 [Rev. E].

[11] Analog Devices. (2023). *AD8099 Tools & Simulations* [Online]. Available: https://www.analog.com/en/products/ad8099.html#product-tools

Long-term Perimeter Protection System Monitoring Using State of a Polarization Analyzer

1st Adrian Tomasov
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
tomasov@vutbr.cz

2nd Petr Dejdar
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
Petr.Dejdar@vut.cz

3th Pavel Záviška
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
Pavel.Zaviska@vut.cz

4th Vladimir Spurny
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
xspurn07@vutbr.cz

5rd Petr Munster
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
munster@vut.cz

6th Tomas Horvath
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
horvath@vut.cz

7th Ondrej Klicnik
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
xklicn02@vutbr.cz

***Abstract*—The paper proposes the deployment of a perimeter and data protection system with long-term data storage in an InfluxDB time-series database enhanced by a graphical interface based on a Grafana dashboard. The perimeter is protected using a fiber optic cable, and the transmitted light is analyzed by a state of polarization analyzer detecting mechanical vibrations. Using a simple spectrum comparison method and a threshold, it can detect a perimeter breach and generate an alert.**

***Index Terms*—information security, optical polarization, perimeter protection, vibration detection**

I. Introduction

Fiber optics technology gathers many advances year by year and indeed becoming the sole communication solution for wide area networks (WAN) or local area networks (LAN), and distribution networks (e.g., passive optical networks), mainly due to their transmission throughput and achievable distance. The considerable benefit of the fiber optic networks is allowing updating the optical transceivers for the new faster ones, while the optical paths remain the same. Thanks to a dense wavelength division multiplexing (DWDM) [1] or flexi–grid in elastic optical networks (EON) [2], many optical transceivers can utilize a single optical path and achieve aggregated throughput up to tens of Terabits per second [3].

The fiber itself has more usage than just a transmission medium. There are many point and distributed sensory systems based on the analysis of transmitted light, which may change its physical attributes, such as frequency, phase, or polarization. The current use cases for optical sensors are mechanical or acoustic vibration analysis using Distributed Acoustic Sensing (DAS) [4], temperature sensing with the help of Brillouin Optical Time Domain Reflectometry (BOTDR) [5], or temperature sensing in biomedical applications using state of the art Fiber Bragg gratings (FBG) [6]–[8]. At the same time, there is testing of sensing systems for a large range of applications from perimeter security [9], [10], pipeline industry [11], nuclear facilities [12], biomedical [13] and many other industries.

However, the fiber-optic paths are not resistant to eavesdropping attacks. There are various ways to wiretap transmitted communication, such as V-groove, fiber bending, splitter insertion, FBG, and others [14], [15]. Another possibility is to eavesdrop on human speech near fiber cable by analyzing phase variations of the back-reflected light [16].

There are already various proposed solutions for eavesdropping and abnormal event detection. The quality of the communication channel between optical transceivers can be expressed using an eye diagram, which includes parameters such as jitter, signal-to-noise ratio, or distortion. The abnormal behavior of these parameters may indicate a potential attack on the optical path. Thus, the eye diagram is analyzed by a convolutional neural network and detects possible eavesdropping attempts [17]. Another option is to analyze the network using Optical Time Domain Reflectometer (OTDR) and analyze its output using convolution or recurrent neural networks [18], [19].

The paper proposes a long-term perimeter and data protection monitoring system on a fiber optic path. The optical sensor is based on our ongoing research of the fast state of polarization change analyzer and experimental anomaly detection methods [20], [21]. The sensor can detect mechanical and loud acoustic vibrations based on the polarization change of the transmitted light caused by fiber manipulation by an attacker or an unauthorized person entering a protected perimeter [22]. This output is captured and analyzed using Raspberry Pi (RPi), and the results are sent into a database running in a cloud environment. The deployment also contains a graphical user interface for data visualization, allowing an

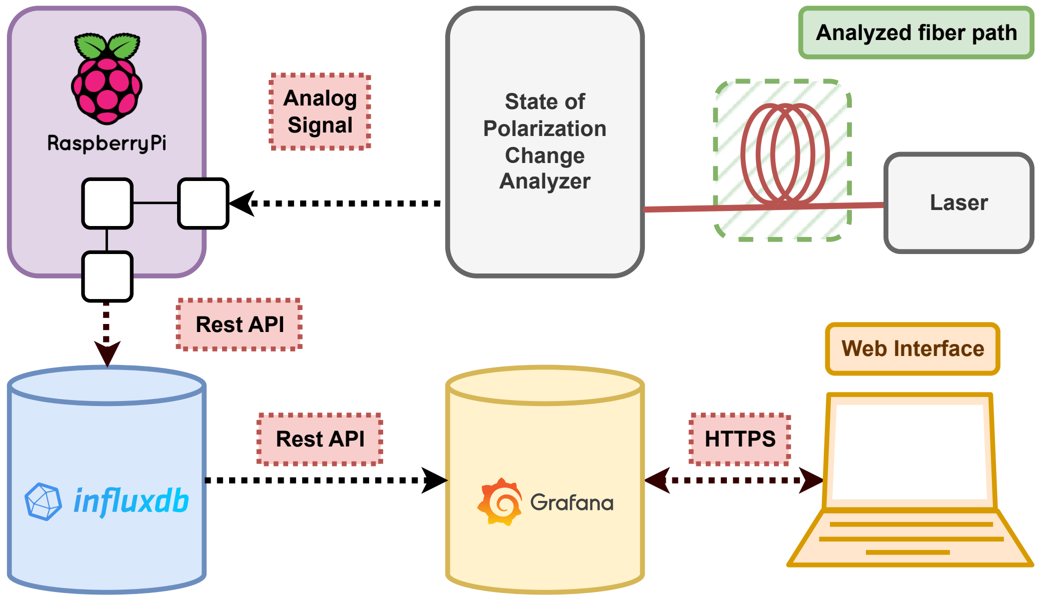


Fig. 1. The proposed deployment of perimeter protection system with long-term data storage consisting of free-to-use open source projects.

anomaly threshold modification and set email alerts if the perimeter is breached or data security is potentially violated.

II. Monitoring System Proposal

The proposed system consists of an optical sensory system with data acquisition and software processing parts, running in an optional environment such as a cloud, bare metal server, or sensor controller.

The hardware part of the perimeter protection system is based on a cost-effective solution analyzing a state of polarization change of transmitted light through the protected fiber path. It can detect mechanical vibrations, including walking on a protected perimeter, fiber manipulation, and others [20]. All vibrations cause tiny differences in the fiber structure and consequently change the transmitted light polarization due to birefringence [23]. The analyzer diagram is shown in Figure 2, where the main components are a polarization beam splitter, and a Koheron LPD 100 balanced photo-detector [24] with a built-in 1550 nm laser. The analyzed fiber segment consists of a common patch cord with G.652D single-mode fiber.

The output is sampled using a Raspberry Pi with an audio card, and the processed data are stored in a time-series database, as shown in Figure 1. All of these components are further described in the following subsections.

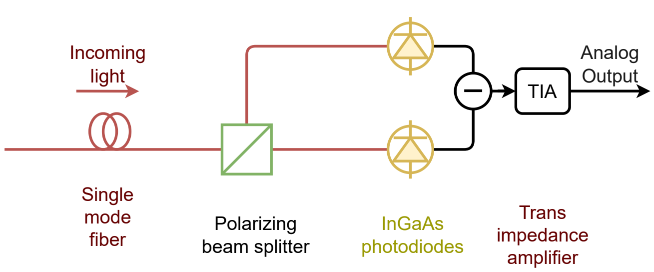


Fig. 2. Schematic diagram of the state of polarization change analyzer, which compares light intensity in orthogonal polarization planes.

A. Signal processing

The analog signal is sampled by a USB audio card connected to RPi [25]. It uses a sampling rate of 8000 samples per second, which is high enough to detect requisite frequencies within the range of 20 to 250 Hz. The incoming signal is divided into overlapping segments of length 4096 samples and a stride[1] of 2048 samples. The individual segments are multiplied with a Hamming window of the corresponding length, and converted into a frequency domain using a single-sided real-valued Discrete Fourier Transform. The spectrum is normalized by subtracting its mean value and then analyzed by comparing means of the bottom and the upper portions of the spectrum, which reduces the signal strength by noise. The whole operation is shown in Eq. (1):

$$\mathrm{DIFF}(\mathbf{x}, p) = \frac{1}{L}\sum_{i=1}^{L} x_i - \frac{1}{N-L}\sum_{i=L}^{N} x_i, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^N$ is a magnitude spectrum vector, p is a number within the range 0 to 1 defining a portion of the lower spectrum part, N is the number of samples in a single window, and it holds that $L = N \cdot p$. The output of this function is compared against the threshold to detect potential perimeter breaches. The higher the output, the more intensive is the intensity of vibrations.

Both portion and threshold parameters are considered as hyper-parameters of the function. The optimal values are evaluated using a grid-search algorithm optimizing the highest accuracy in the anomaly detection occurrences. Furthermore,

[1] The stride parameter represents the number of samples by which the window shifts.

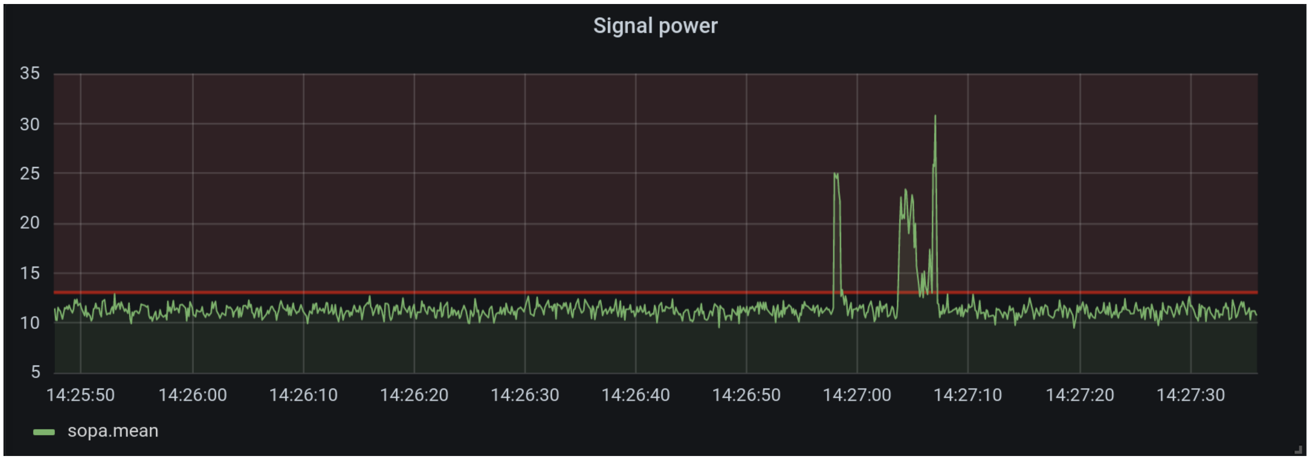


Fig. 3. The Grafana dashboard with the time series of signal processing method output indicating perimeter breach.

the threshold parameter should be fine-tuned for each specific deployment based on location. The optimal values can be found in a testing measurement, where the threshold is set to a maximal achievable value by Eq. (1) or a 99^{th} quantile of the equation output.

B. Long-term monitoring

The output of the signal processing function is sent into a long-term monitoring system consisting of InfluxDB [26] time series database and Grafana [27] system for graphical representation of stored data. InfluxDB is optimized for time series data from many sources by allowing storing data in several separate tables. Furthermore, it allows distinguishing data in the same tables using tags. Our proposed system uses these to differentiate between various deployed perimeter protection systems.

Another key attribute of InfluxDB is the possibility to aggregate/interpolate requested data using mathematical functions (e.g., mean) to reduce data request size, which is very useful, particularly in requests with very long time ranges (e.g., months, years).

All data stored in the database are graphically shown in charts using the Grafana platform. Grafana is an open-source, user-friendly project for real-time charts and similar system monitoring tools. It has a native integration with many data sources, including InfluxDB. Many different charts in Grafana allow the creation of many advanced system monitoring dashboards. This proposal uses a simple time-series chart enhanced with a threshold to decide whether the signal is nominal or abnormal. Furthermore, if the threshold is exceeded, it generates an email alert to a responsible person.

InfluxDB and Grafana are standalone services that can run in a cloud environment or any possible Linux based system. The only requirement is network visibility because they use a representational state transfer application interface (RestAPI) for communication. However, the cloud deployment option is preferred because it can utilize and store data for more sensory systems in one place.

The whole deployment proposal is shown in Figure 1, where ADC stands for analog-digital converter, DSP is digital signal processing, ETH is ethernet, Rest API stands for representational state transfer application interface, and HTTPS abbreviation means hypertext transfer protocol secure.

III. Implemented Results

The core program responsible for sampling, data processing, and writing to the database in RPi is written in python3 with third-party opensource libraries, such as `numpy` for signal processing, `influxdb-client` for database communication, and last but not least, `soundcard` for simple data sampling from the audio card. It automatically starts on system start-up using a SystemD service and runs in an infinite loop.

InfluxDB and Grafana are installed into a virtual machine with an assigned public IP address secured by a firewall. Furthermore, both services use `certbot` [28] program to request a valid certificate inside the public key infrastructure (PKI), which allows using trusted connections between them and users. It creates a trustworthy channel for the following password authentication.

The hardware part of the proposed system is deployed in a test laboratory with fiber optics cable installed in ceilings. It should detect anyone opening the door and entering the room. Figure 4 shows the fiber path and the laboratory itself. Figure 3 shows a time-series chart from the Grafana dashboard containing the difference of spectrum means, as is defined in Equation (1), represented by a green line and threshold line drawn by a red color with the abnormal red highlighted region. There are three visible abnormal events, the first is door closure, and the second is fiber manipulation. All these events generated alerts and sent them to a desired e-mail address.

Furthermore, a specific tag is added to a database write request to distinguish the data from this specific instance.

IV. Conclusion

The paper describes deploying a cost-effective perimeter and fiber protection system with all required monitoring software

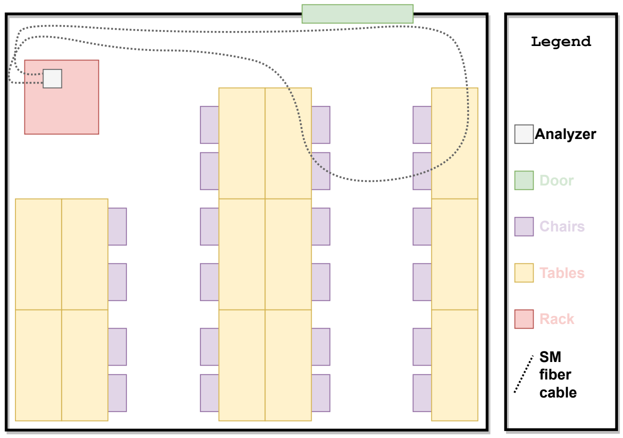


Fig. 4. Analyzed laboratory with optical fiber in ceilings shown by a dotted curved line.

in a real-life scenario. The hardware part is based on the state of the polarization analyzer and a Raspberry Pi with an audio card for analog data sampling and processing. The potential security breach in the protected perimeter is detected using our ongoing research's simple spectrum comparison method. The results are stored in the InfluxDB time-series database for long-term monitoring. These data are displayed by line charts using the Grafana web application, which compares the data against a defined threshold. Additionally, the application can send an e-mail alert if the output, if comparison, exceeds the defined threshold.

Future work will focus on adapting this deployment proposal to classification problems based on neural networks. They require more computational power to run a neural network model; thus, the RPi is unsuitable. Also, their output is more complex, which is a vector of probabilities for each classification class. Another critical part is the graphical representation of these vectors together with class counters and similar.

ACKNOWLEDGMENT

Research described in this paper was supported by the Ministry of the Interior of the Czech Republic, program IMPAKT1, under grant VJ01010035, project "Security risks of photonic communication networks".

REFERENCES

[1] C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *IEEE Journal on Selected areas in Communications*, vol. 8, no. 6, pp. 948–964, 1990.

[2] V. López, L. Velasco, *et al.*, "Elastic optical networks," *Architectures, Technologies, and Control, Switzerland: Springer Int. Publishing*, 2016.

[3] N. A. Shevchenko, S. Nallaperuma, and S. J. Savory, "Maximizing the information throughput of ultra-wideband fiber-optic communication systems," *Opt. Express*, vol. 30, pp. 19320–19331, May 2022.

[4] Z. He and Q. Liu, "Optical fiber distributed acoustic sensors: A review," *J. Lightwave Technol.*, vol. 39, pp. 3671–3686, Jun 2021.

[5] Q. Bai, Q. Wang, D. Wang, Y. Wang, Y. Gao, H. Zhang, M. Zhang, and B. Jin, "Recent advances in brillouin optical time domain reflectometry," *Sensors*, vol. 19, no. 8, p. 1862, 2019.

[6] P. Roriz, S. Silva, O. Frazao, and S. Novais, "Optical fiber temperature sensors and their biomedical applications," *Sensors*, vol. 20, no. 7, p. 2113, 2020.

[7] D. Tosi, S. Poeggel, I. Iordachita, and E. Schena, "Fiber optic sensors for biomedical applications," in *Opto-mechanical fiber optic sensors*, pp. 301–333, Elsevier, 2018.

[8] G. Palumbo, D. Tosi, A. Iadicicco, and S. Campopiano, "Analysis and design of chirped fiber bragg grating for temperature sensing for possible biomedical applications," *IEEE Photonics Journal*, vol. 10, no. 3, pp. 1–15, 2018.

[9] K. Witas, M. Stolarik, M. Pinka, S. Zabka, J. Jalowiczor, M. Novak, and R. Jaros, "Perimetric monitoring: A comparison of a classical seismic sensor and fiber-optic interferometric sensor," in *Optical Sensors 2019*, vol. 11028, p. 110282K, International Society for Optics and Photonics, SPIE, 2019.

[10] F. Liu, H. Zhang, X. Li, Z. Li, and H. Wang, "Intrusion identification using gmm-hmm for perimeter monitoring based on ultra-weak fbg arrays," *Opt. Express*, vol. 30, pp. 17307–17320, May 2022.

[11] W. Gan, N. Liu, Y. Wang, Y. Zhang, X. Chen, J. Tang, and C. Zhang, "Two-parameter method for identification and location of leaks based on weak fbg for steam pipelines," *Optical Fiber Technology*, vol. 74, p. 103095, 2022.

[12] J. Wu, M. Wang, K. Zhao, S. Huang, M. A. S. Zaghloul, R. Cao, D. Carpenter, G. Zheng, S. D. Rountree, and K. P. Chen, "Distributed fiber sensors with high spatial resolution in extreme radiation environments in nuclear reactor cores," *Journal of Lightwave Technology*, vol. 39, no. 14, pp. 4873–4883, 2021.

[13] P. Roriz, S. Silva, O. Frazão, and S. Novais, "Optical fiber temperature sensors and their biomedical applications," *Sensors*, vol. 20, no. 7, 2020.

[14] V. Spurny, P. Munster, A. Tomasov, T. Horvath, and E. Skaljo, "Physical layer components security risks in optical fiber infrastructures," *Sensors*, vol. 22, no. 2, p. 588, 2022.

[15] H. Si, H. Liu, and H. Ma, "Optical fiber communication network eavesdropping and defensive measures," in *Proceedings of the 2nd International Forum on Management, Education and Information Technology Application (IFMEITA 2017)*, pp. 307–310, Atlantis Press, 2018/02.

[16] H. Hao, Z. Pang, G. Wang, and B. Wang, "Indoor optical fiber eavesdropping approach and its avoidance," *Opt. Express*, vol. 30, pp. 36774–36782, Sep 2022.

[17] M. Liu, Y. Li, H. Song, Z. Tu, Y. Zhao, and J. Zhang, "Experimental demonstration of optical fiber eavesdropping detection based on deep learning," in *Asia Communications and Photonics Conference*, pp. T1H–4, Optical Society of America, 2019.

[18] H. Song, Y. Li, M. Liu, K. Wang, J. Li, M. Zhang, Y. Zhao, and J. Zhang, "Experimental study of machine-learning-based detection and location of eavesdropping in end-to-end optical fiber communications," *Optical Fiber Technology*, vol. 68, p. 102669, 2022.

[19] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, "Machine-learning-based anomaly detection in optical fiber monitoring," *J. Opt. Commun. Netw.*, vol. 14, pp. 365–375, May 2022.

[20] A. Tomasov, P. Dejdar, T. Horvath, and P. Munster, "Physical fiber security by the state of polarization change detection," in *Fiber Optic Sensors and Applications XVIII*, vol. 12105, pp. 52–56, SPIE, 2022.

[21] M. Ruzicka, L. Jabloncik, P. Dejdar, A. Tomasov, V. Spurny, and P. Munster, "Classification of events violating the safety of physical layers in fiber-optic network infrastructures," *Sensors*, vol. 22, no. 23, 2022.

[22] P. Dejdar, V. Myska, P. Munster, and R. Burget, "Trains detection using state of polarization changes measurement and convolutional neural networks," in *2021 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, pp. 1–4, 2021.

[23] A. Smith, "Birefringence induced by bends and twists in single-mode optical fiber," *Applied optics*, vol. 19, no. 15, pp. 2606–2611, 1980.

[24] Y. M. Klimkov, M. Shribak, and E. Yur'eva, "A balance method of measuring birefringence for normal reflection," *Measurement Techniques*, vol. 36, no. 7, pp. 759–762, 1993.

[25] B. Horan, *Practical Raspberry Pi*. Apress, 2013.

[26] M. Nasar and M. A. Kausar, "Suitability of influxdb database for iot applications," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 10, pp. 1850–1857, 2019.

[27] M. Chakraborty and A. P. Kundan, "Grafana," in *Monitoring Cloud-Native Applications: Lead Agile Operations Confidently Using Open Source Software*, pp. 187–240, Springer, 2021.

[28] A. Mai, O. Schedler, E. Weippl, and K. Krombholz, "Are https configurations still a challenge?: Validating theories of administrators' difficulties with tls configurations," in *HCI for Cybersecurity, Privacy and Trust: 4th International Conference, HCI-CPT 2022*, pp. 173–193, Springer, 2022.

Review of In-situ and Ex-situ Techniques for Characterization of Li-ion Batteries

1st Ondřej Klvač
Department of Electrical and Electronic Technology
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xklvac02@vutbr.cz

2nd Tomáš Kazda
Department of Electrical and Electronic Technology
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
kazda@vutbr.cz

***Abstract*—This paper describes the analytical techniques and sample preparation methods used in in-situ and ex-situ investigations of Li-ion batteries (LIBs). The introduction briefly describes the current state of Li-ion batteries and their aging mechanisms. The article is further divided into a description of ex-situ and in-situ techniques. Ex-situ techniques include a detailed description of sample preparation for repeatable results using Broad Ion Beam (BIB) polisher. In-situ techniques are described in the form of a report of selected published experiments. The paper is further extended with our experiments involving ex-situ analyses of cylindrical and flat cell types using BIB in planar mode and cross-section mode. In the case of in-situ experiments, the design of the cell inside the SEM chamber is described. Two experiments have been performed - on the surface of a MEMS chip, and also using a micromanipulator as a current collector of the investigated electrode.**

***Index Terms*—Li-ion, SEM, FIB, EDS, BIB, battery, energy storage, in-situ, ex-situ, sample preparation, ionic liquid**

I. Introduction

Depending on the application, we require from electrochemical current sources high energy density (gravimetric, volumetric), long lifetime (number of possible charges - cycle lifetime and service life), good resistance to external conditions (e.g. temperature), resistance to high loads (charging and discharging currents), low cost, environmental compatibility, safety, and others. Lithium-ion batteries (LIBs) meet these requirements for most of today's applications, making them the most widely used battery type. However, as time goes on, the demands on batteries are increasing. In the future, it is also necessary to take into account the limited supply of raw materials for battery production (mainly nickel and cobalt). Therefore, there is an effort to find new materials to respond to these challenges [1] [2] [3].

In the future, the use of materials such as sulfur or silicon, high-voltage spinels, and solid electrolytes are being considered. These technologies have been known for a long time, but have not been used commercially, mainly due to the instability of such a system and the very low cycle life. The aim is to understand the degradation mechanisms and propose solutions. In the manufacturing process, the aim is to prevent degradation and potentially hazardous conditions [3] [4] [5].

Since LIBs are secondary cells that can be recharged repeatedly, the reactions taking place during cycling are ideally 100 % reversible. In practice this is not the case, certain changes are irreversible, causing aging and irreversible loss of capacity and performance. This may occur not only during operation but also by improper storage in an inactive state [6].

The main mechanisms include aging due to phase transformations of the crystallographic structure of the electroactive materials, electrolyte decomposition, and Solid Electrolyte Interphase (SEI) formation. When the SEI layer is disrupted, a new layer is formed, which further consumes lithium, and lithium dendrite formation can also occur on the anode side. Transition metals may also dissolve from the cathode and migrate toward the anode. A number of techniques are available to analyze these processes. The most versatile is the use of scanning electron microscopy together with associated methods such as EDS [7].

This work continues with a description of sample preparation and an overview of the ex-situ and in-situ techniques used using SEM to characterize Li-ion battery systems. The individual parts are complemented by experience from our experiments.

II. Ex-situ techniques

For comparable results, the battery must be analyzed under similar defined conditions. A common reference is the open-circuit voltage (OCV) and the depth of charge/discharge (SOC/SOD) derived from it. Since the OCV is also influenced by other factors (previous load), it is advisable to hold the defined voltage for several hours or to perform a certain number of identical cycles for all analyzed samples. From a safety point of view, it is recommended to discharge the battery as much as possible. On the other hand, the voltage should not leave the nominal values specified in the datasheet in order to avoid material changes not caused by cycling (e.g. dissolution of current collectors) [8] [9].

The battery contains materials highly sensitive to contact with air humidity, oxygen, and nitrogen. When exposed to air, the samples are visibly degraded within seconds. In addition, the frequently used lithium salt $LiPF_6$ reacts with air moisture to form HF. Therefore, it is necessary to open cells in a glovebox filled with inert gas with controlled levels of O_2 and

DOI 10.13164/eeict.2023.207

H_2O (ideally also N_2) or dry laboratory environments with ventilation [8] [10] [11].

It is preferable to avoid short-circuiting during disassembly. Therefore, non-conductive tools should be used and excessive pressure should be avoided. The standard procedure is to remove the housing and then separate the electrodes. In the case of whole battery modules in electric vehicles, the interconnected flat-type cells can be separated by using a nylon string. However, excessive mechanical stress must be avoided. Before disassembling an unknown cell type, the X-ray method can be used to determine the optimal location for the cut [10] [12].

Inspection of multilayered structures in the cross-section can be problematic. For cylindrical cells, a metal casing can be used as a fixation to hold the layers together during mechanical polishing. However, the pouch and prismatic cells need to be fixed with, for example, a clamp or embedded in epoxy resin [12].

Further sample preparation is often required before analysis. Grinding and polishing are common preparation techniques to remove roughness and obtain a smooth surface. However, these traditional mechanical techniques apply shear forces to the sample that may result in damage to the sample. The electrode layers may delaminate and grooves can be created on the surface. Broad ion beam (BIB) technology is therefore often used for final polishing. These systems sputter the sample material with noble gas ions (typically argon). No mechanical stress is applied, leaving extremely flat surfaces and cuts with minimal artifacts. BIB is typically used in two different modes - surface (planar) polishing and cross-section polishing. During polishing, the sample is heated, so it is necessary to ensure sufficient heat dissipation or to use cooling (typically a Peltier cell or liquid nitrogen) [12] [13] [14] [15].

In the case of surface polishing, the beam hits directly the rotating specimen. The surface is mechanically pre-polished and then only a few micrometers of material are sputtered. Figure 1 shows an example of a sample prepared in this way. It is a 18650 cylindrical Li-ion cell. It was discharged to 0.1 V, cut perpendicularly with a hacksaw, and mechanically pre-polished on SiC grinder paper. Isopropyl alcohol was used as the grinding medium. This was followed by final polishing using a two-ion gun BIB polisher Fishione Model 1061 Ion Mill with Cryo Cooling. SEM inspection of the prepared area provides information about the internal arrangement of the electrodes including defects on the current collector of the anode (red arrows) [16].

Thin materials up to 1 mm are suitable for cross-sections. The sample is placed behind the mask (hard-to-sputter material, e.g. titanium) in a slight overhang (overlap of the sample above the mask surface, usually tens of microns). The overlapping part is sputtered off and a sharp edge is created. The size of the overhang is a compromise between time and the certainty of removing the part containing the defects caused by shearing/cutting the material. The specimen is rotated alternately during the process (so-called rocking or oscillation), which ensures that the beam is always coming

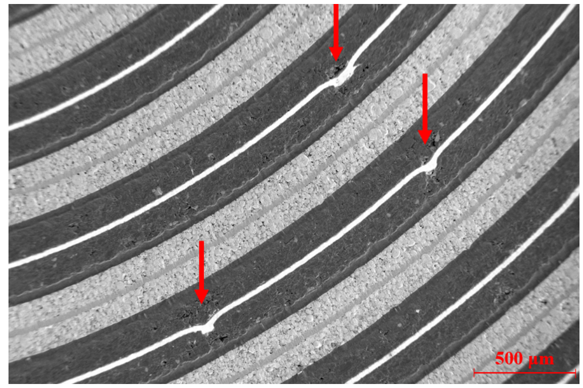


Fig. 1. Cross-section of whole 18650 cyllindrical cell prepared using BIB in planar mode.

from a different angle. This enlarges the polished area and reduces artifacts.

Typically, individual electrodes or a separator are prepared in this way. Figure 2 is a comparison of the same area of the cathode before polishing (only cut with scissors) and after polishing. A BIB polisher Fishione Model 1061 Ion Mill was used for preparation in cross-section mode with an overhang of approx. 20 µm and an oscillation of ±30°.

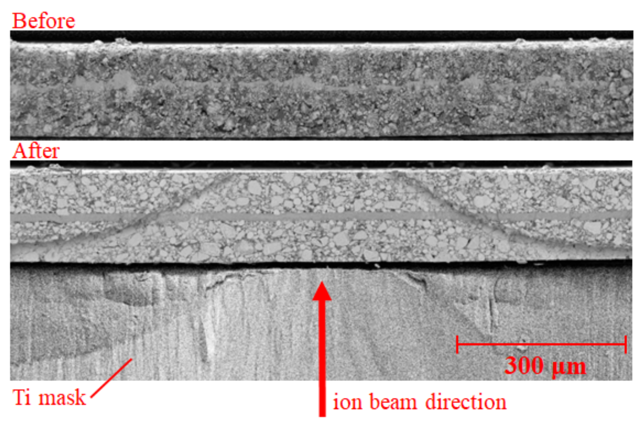


Fig. 2. Cross-section of battery cathode prepared using BIB. Comparison of the surface before and after polishing.

After shearing, the structure is damaged and only the current collector is hardly recognizable. After polishing, a region with a Gaussian profile (due to the beam intensity) is apparent where particles of electroactive material, binder, and current collector are recognizable. It can be determined that this is a blend of two electroactive materials, there is no significant grain cracking or delamination, the electrode has no surface coating (e.g. alumina layer) and there is no significant SEI layer.

III. In-situ techniques

One of the first in-situ SEM experiments was published in 1988. The cell was prepared by deposition of thin layers of active material (FeS, TiS_2, V_6O_{13}) on current collectors connected through a vacuum feedthrough outside the microscope. A solid-state electrolyte (polyether with ethylene oxide base and $LiClO_4$ salt) was used. Lithium metal was used as a counter electrode. The cell was assembled in a glove box and transported in a short time (about 30 s) to the SEM to reduce degradation in air. The cell was scanned from the side of the sandwich structure and was heated during cycling due to poor conductivity. A schematic drawing of the cell is shown in Fig. 3. No changes in the structure could be observed for the stable TiS_2, V_6O_{13} showed small cracks, and FeS completely decayed [17].

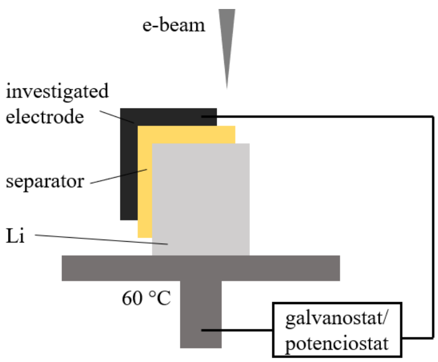


Fig. 3. Schematic drawing of the in-situ experiment – cell investigated from the side view.

Another approach is to observe from the top. In 2011, a paper was published in which the half-cell consisted of lithium metal placed on a stage connected to a microscope frame, with an ionic liquid in a separator (TFSI + LiTFSI) and SnO_2 pressed into a steel grid as the active material. The lithium and separator with electrolyte were transported from the glovebox in a hermetically insulated capsule, which was opened after pumping the SEM chamber by rotating the stage. The grid with the active material was then attached to the separator. The disadvantage of the arrangement is that the electrode is examined from the top, on the side opposite to the lithium. A schematic drawing of the cell is shown in Fig. 4. However, it was possible to observe the nanoscale structural changes in real time. Also, local charging and deposition of lithium were observed during scanning at locations far from the metal lattice (interaction with the electron beam) [18].

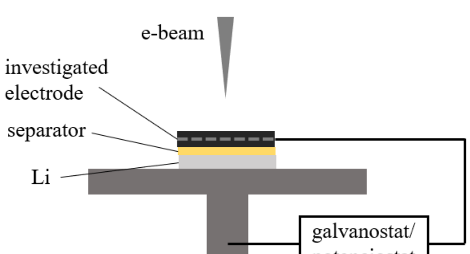


Fig. 4. Schematic drawing of the in-situ experiment – cell investigated from the top view.

Deposition of the electroactive material onto the grid can also be done electrophoretically without the need for a binder. The 2016 work thus creates an electrode from silicon microplates in which lithium concentration could be observed by varying the contrast of the BSE image [19].

Another way of investigation is to fabricate the cell directly in the SEM chamber using FIB-SEM technology and a micromanipulation needle as a current collector (see Fig. 5 for a schematic drawing). A 2013 scientific paper describes an LTO/ionic liquid/NCA battery. A small piece of the NCA-containing electrode material was cut out using a focused ion beam (FIB), conductively and mechanically joined by assisted deposition with a micromanipulation needle, and dipped in ionic liquid. An inspection hole was cut in the middle of the electrode to observe changes during cycling [20].

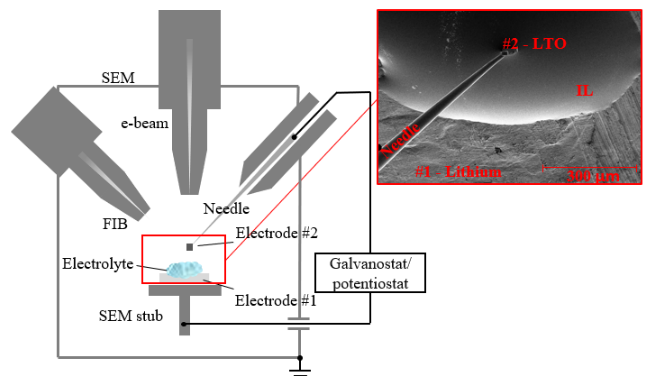


Fig. 5. Schematic drawing of the system using a micromanipulator as a current collector (left); SEM image taken during the experiment (right).

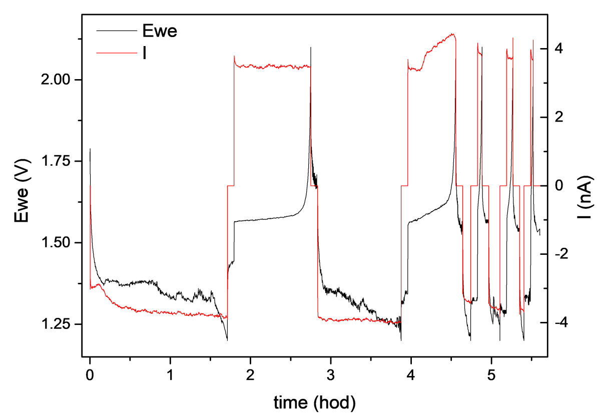


Fig. 6. Electrochemical cycling of the cell where a micromanipulator was used as a current collector. The graph shows that the capacitance decreases with each cycle.

We used this method for the construction of a lithium/ionic liquid/lithium titanate (LTO) half-cell [21]. For the experiment, we used a Scios 2 SEM (Thermo Fisher Scientific) equipped with an EasyLift micromanipulator, a gallium FIB, and a MultiChem Gas Delivery System. A piece of lithium was placed on the stub, which was electrically connected through a vacuum feedthrough to the potentiostat Biologic SP-150 as a shared reference and counter electrode. The lithium metal was placed on the stub and dropped with 1-Ethyl-3-

methylimidazolium tetrafluoroborate (EMIMBF) ionic liquid in a mixture with $LiBF_6$ salt in 0.5M concentration serving as the electrolyte. A chunk of LTO material was glued to the micromanipulation needle by assisted carbon deposition. The needle was modified, insulated from the microscope frame, and connected to the potentiostat as a working electrode. After dipping the LTO into the electrolyte, several cycles were successfully performed (see Fig. 6). The experiment was strongly affected by electromagnetic interference and side reactions of the lithium metal with the electrolyte. We are currently working on improvements.

The vacuum environment of SEM is convenient for the study of dendrite growth. The sample can be observed on the reaction side as a function of time without the necessity of opening the cell after cycling. Published results confirmed lithium deposition and dendrite growth at increased rates at higher currents for polymer and liquid electrolytes. In the case of sulfur-based solid-state cells, dendrite growth was also observed, although this was not predicted. However, the fact is that higher current densities can cause cracks in the solid electrolyte and hence subsequent inhomogeneous deposition of lithium [22] [23] [24].

To study dendrite growth, a special cell consisting of a pair of contacts on a wafer with deposited lithium, connected to a current source, was published in 2017 (see Fig. 7). This wafer is then attached to another wafer with a SiN_x inspection window transparent to electrons, preventing the evaporation of volatile electrolytes in a vacuum. A constant current flows through the cell and the interface between the lithium metal and the electrolyte was monitored. The effect of additives on enhancement or suppression of dendrite growth was described [28].

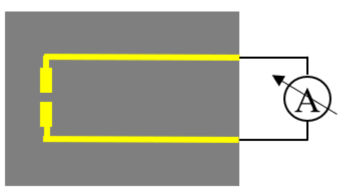


Fig. 7. Cell for the study of dendrite growth.

We used a similar setup for our first in-situ experiment [25]. The cell construction was performed on the surface of a MEMS chip (see Figure 8) [26]. This chip was designed for heating and biasing experiments in a SEM chamber (Thermo Fisher Scientific) [27]. The biasing contacts were used as current collectors for electroactive materials, directly connected to the potentiostat Biologic SP-150 via vacuum feedthrough. Chunks of approximately 30 x 50 x 20 µm were cut from the electroactive materials (metallic lithium and LTO) using FIB on Helios Hydra 5 UX DualBeam and Helios 5 UC DualBeam SEMs (Thermo Fisher Scientific). By assisted carbon deposition using the MultiChem Gas Delivery system, they were glued to an EasyLift micromanipulation needle, transferred and aligned with biasing contacts, and connected by assisted carbon deposition as well. EMIMBF ionic liquid in a mixture with $LiBF_6$ salt in 0.5M concentration was used as an electrolyte, which was transferred to the area between the electrodes after their construction.

The system is very sensitive to surface contamination around the electrodes. Cell capacitance is at most tens of nAh, so even a very thin layer of redeposited material can short-circuit the battery. Care must also be taken with the amount of used electrolyte. Once the electrodes are flooded, it is no longer possible to disconnect the cell and if the electrolyte covers the investigated area, it cannot be removed and there is no possibility of further investigating this area. On the other hand, the preparation process can be well automated and have a high repeatability of results. Also, the preparation process can be managed in the SEM chamber without the need for air transfer. The heating area of the chip can also be used for the synthesis of electroactive materials which will be subsequently used for electrode preparation. We are currently dealing with this concept [25].

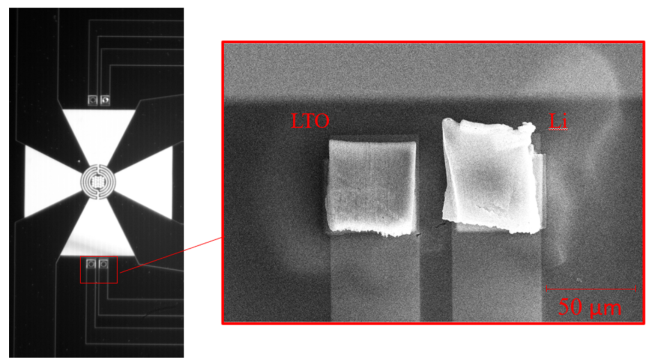


Fig. 8. MEMS chip used for the cell construction (left); Detail of the biasing contacts area with the electroactive material (right).

IV. Conclusion

Scanning electron microscopy provides high-resolution images of LIBs surface or cross-sections and is a powerful method for the study of most of the important processes taking place inside Li-ion batteries. It can be used both in adjusting the manufacturing process in the factory as well as in research in the laboratory.

However, the analysis has its specificities because LIBs contain sensitive materials. The vacuum environment of the SEM can limit the observation in their native state. Due to the low mechanical stability, the samples must be handled with care to distinguish preparation-induced defects from cycling-induced defects. Therefore, BIB technology is often used for sample preparation, especially for ex-situ analysis.

Although SEM alone provides detailed information on morphology and microstructure, it does not provide information about the associated electrochemical processes. For these purposes, in-situ experiments that allow SEM analysis and electrochemical cycling at the same time can be beneficial.

ACKNOWLEDGMENT

This work was supported by the specific graduate research of the Brno University of Technology No. FEKT-S-23-8286. The work was developed in cooperation with Thermo Fisher Scientific Brno.

REFERENCES

[1] Y. Liang, C. -Z. Zhao, H. Yuan, Y. Chen, W. Zhang, J. -Q. Huang, D. Yu, Y. Liu, M. -M. Titirici, Y. -L. Chueh, H. Yu, and Q. Zhang, "A review of rechargeable batteries for portable electronic devices", *InfoMat*, vol. 1, no. 1, pp. 6-32, 2019.

[2] Y. Tao, C. D. Rahn, L. A. Archer, and F. You, "Second life and recycling: Energy and environmental sustainability perspectives for high-performance lithium-ion batteries", *Science Advances*, vol. 7, no. 45, Nov. 2021.

[3] H. -J. Kim, T. N. V. Krishna, K. Zeb, V. Rajangam, C. V. V. M. Gopi, S. Sambasivam, K. V. G. Raghavendra, and I. M. Obaidat, "A Comprehensive Review of Li-Ion Battery Materials and Their Recycling Techniques", *Electronics*, vol. 9, no. 7, 2020.

[4] G. Q. Liu, L. Wen, and Y. M. Liu, "Spinel LiNi0.5Mn1.5O4 and its derivatives as cathodes for high-voltage Li-ion batteries", *Journal of Solid State Electrochemistry*, vol. 14, no. 12, pp. 2191-2202, 2010.

[5] M. Rana, S. A. Ahad, M. Li, B. Luo, L. Wang, I. Gentle, and R. Knibbe, "Review on areal capacities and long-term cycling performances of lithium sulfur battery at high sulfur loading", *Energy Storage Materials*, vol. 18, pp. 289-310, 2019.

[6] C. D. Rahn, *Battery systems engineering*. Chichester, West Sussex, United Kingdom: John Wiley & Sons Ltd., Publication, 2013, p. 16-44.

[7] C. Lin, A. Tang, H. Mu, W. Wang, and C. Wang, "Aging Mechanisms of Electrode Materials in Lithium-Ion Batteries for Electric Vehicles", *Journal of Chemistry*, vol. 2015, pp. 1-11, 2015.

[8] T. Waldmann et al., "Review—Post-Mortem Analysis of Aged Lithium-Ion Batteries: Disassembly Methodology and Physico-Chemical Analysis Techniques", *Journal of The Electrochemical Society*, vol. 163, no. 10, pp. A2149-A2164, Sep. 2016.

[9] Y. Kobayashi, T. Kobayashi, K. Shono, Y. Ohno, Y. Mita, and H. Miyashiro, "Decrease in Capacity in Mn-Based/Graphite Commercial Lithium-Ion Batteries", *Journal of The Electrochemical Society*, vol. 160, no. 8, pp. A1181-A1186, May 2013.

[10] N. Williard, B. Sood, M. Osterman, and M. Pecht, "Disassembly methodology for conducting failure analysis on lithium–ion batteries", *Journal of Materials Science: Materials in Electronics*, vol. 22, no. 10, pp. 1616-1630, 2011.

[11] T. Waldmann, S. Gorse, T. Samtleben, G. Schneider, V. Knoblauch, and M. Wohlfahrt-Mehrens, "A Mechanical Aging Mechanism in Lithium-Ion Batteries", *Journal of The Electrochemical Society*, vol. 161, no. 10, pp. A1742-A1747, Jul. 2014.

[12] Kovachev et al., "Analytical Dissection of an Automotive Li-Ion Pouch Cell", *Batteries*, vol. 5, no. 4, 2019.

[13] M. Indrikova, S. Grunwald, F. Golks, A. Netz, B. Westphal, and A. Kwade, "The Morphology of Battery Electrodes with the Focus of the Conductive Additives Paths", *Journal of The Electrochemical Society*, vol. 162, no. 10, pp. A2021-A2025, Jul. 2015.

[14] A. Sakuda, A. Hayashi, and M. Tatsumisago, "Sulfide Solid Electrolyte with Favorable Mechanical Property for All-Solid-State Lithium Battery", *Scientific Reports*, vol. 3, no. 1, Dec. 2013.

[15] Y. Shimauchi, S. Ikemoto, S. Ohmori, and T. Itoi, "IEx Situ/i Electron Microscopy Study of the Lithiation of Single-Crystal Si Negative Electrodes during Charge Reaction in a Lithium–Ion Battery", *MATERIALS TRANSACTIONS*, vol. 60, no. 11, pp. 2328-2335, 2019.

[16] P. Blažek et al., "Multiscale 3D Analysis of Defects and Temporal Development of Electrode Morphology in Lithium-Ion Batteries By X-Ray Computed Tomography", *ECS Transactions*, vol. 105, no. 1, pp. 69-76, Nov. 2021.

[17] P. BAUDRY, "In situ observation by SEM of positive composite electrodes during discharge of polymer lithium batteries", *Solid State Ionics*, vol. 28-30, pp. 1567-1571, 1988.

[18] D. Chen, S. Indris, M. Schulz, B. Gamer, and R. Mönig, "In situ scanning electron microscopy on lithium-ion battery electrodes using an ionic liquid", *Journal of Power Sources*, vol. 196, no. 15, pp. 6382-6387, 2011.

[19] C. -Y. Chen et al., "In situ Scanning Electron Microscopy of Silicon Anode Reactions in Lithium-Ion Batteries during Charge/Discharge Processes", *Scientific Reports*, vol. 6, no. 1, Dec. 2016.

[20] D. J. Miller, C. Proff, J. G. Wen, D. P. Abraham, and J. Bareño, "Observation of Microstructural Evolution in Li Battery Cathode Oxide Particles by In Situ Electron Microscopy", *Advanced Energy Materials*, vol. 3, no. 8, pp. 1098-1103, 2013.

[21] "Proceedings I of the 28th Conference STUDENT EEICT 2022", in *Proceedings I of the 28th Conference STUDENT EEICT 2022*, 2022, pp. 398-402.

[22] M. Dollé, L. Sannier, B. Beaudoin, M. Trentin, and J. -M. Tarascon, "Live Scanning Electron Microscope Observations of Dendritic Growth in Lithium/Polymer Cells", *Electrochemical and Solid-State Letters*, vol. 5, no. 12, 2002.

[23] F. Orsini et al., "In situ Scanning Electron Microscopy (SEM) observation of interfaces within plastic lithium batteries", *Journal of Power Sources*, vol. 76, no. 1, pp. 19-29, 1998.

[24] M. Nagao, A. Hayashi, M. Tatsumisago, T. Kanetsuku, T. Tsuda, and S. Kuwabata, "In situ SEM study of a lithium deposition and dissolution mechanism in a bulk-type solid-state cell with a Li2S–P2S5 solid electrolyte", *Physical Chemistry Chemical Physics*, vol. 15, no. 42, 2013.

[25] O. Klvac, T. Kazda, O. Cech, Y. Fam, and L. Novak, "Electrochemical Cell Preparation on MEMS Chip Surface Inside Scanning Electron Microscope", *ECS Transactions*, vol. 105, no. 1, pp. 655-663, Nov. 2021.

[26] L. Mele et al., "A MEMS-based heating holder for the direct imaging of simultaneous in-situ heating and biasing experiments in scanning/transmission electron microscopes", *Microscopy Research and Technique*, vol. 79, no. 4, pp. 239-250, 2016.

[27] L. Novák, J. Stárek, T. Vystavěl, and L. Mele, "MEMS-based Heating Element for in-situ Dynamical Experiments on FIB/SEM Systems", *Microscopy and Microanalysis*, vol. 22, no. S3, pp. 184-185, 2016.

[28] G. Rong et al., "Liquid-Phase Electrochemical Scanning Electron Microscopy for In Situ Investigation of Lithium Dendrite Growth and Dissolution", *Advanced Materials*, vol. 29, no. 13, 2017.

Integrating Real-Environment 4G and 5G Mobile Data Patterns into Network Simulations for Delay Analysis

1st Aneta Kolackova
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
aneta.kolackova@vut.cz

2nd Jan Jerabek
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
jerabekj@vut.cz

3nd Jiri Hosek
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
hosek@vut.cz

***Abstract*—This article presents a simulation testbed for analyzing a mobile transport network using open-source networking tools and protocols. The topology was designed according to the 5G transport network architecture, with Segment Routing for IPv6 (SRv6) technology and The Fast Data Project Vector Packet Processor (FD.io VPP) utilized for packet processing. The Trex traffic generator was used to simulate data transmission, and the Docker platform enabled centralized access and management of the topology. Real data patterns collected from one cell site was inserted into the created topology to create output for future research. The created testbed provides a valuable tool for analyzing the mobile transport network's performance and future optimizing based on needs of network slices.**

***Index Terms*—SRv6, network slicing, FD.io VPP, real data set, testbed, transport network**

I. Introduction

Mobile networks have become an integral part of modern communication systems, with millions of users relying on them for voice, video, data, and internet services. A mobile transport network is a key component of such networks, responsible for transporting mobile traffic from end-user devices to various network services. It connects the radio access network (RAN) and core network elements, facilitating the transmission of data and voice traffic between them. A typical mobile transport network comprises wired and wireless links such as fiber optic cables, microwave links, and satellite links and also routers and switches.

As the world becomes increasingly reliant on wireless communication, the development of 5G technology has become a key priority for many industries. One of the key challenges of 5G technology is the design of a mobile transport network that can support the high-bandwidth, low-latency data transfers required by 5G applications. To address this challenge, researchers have been utilizing the network simulation as a way to test and optimize different network topologies and configurations. Researchers also have explored the integration of real-environment data into the topology of these networks, and the use of network segment routing technology (SR), specifically SRv6, which is seen as a potential solution enabling network slicing in IP-based 5G systems [7].

Network slicing technology plays a key role in this paper because by using SRv6 we are able to break down individual slices and treat them differently. Using real network data, we are then able to optimize the network and obtain data for further testing [1].

In this text, we explore the use of network simulation in the context of mobile 5G transport networks. We begin by describing related research activities, including efforts to integrate real-world environment data into network simulations. We note that the use of network simulation with real mobile data in transport networks is unique. We then focus on the future of network simulation in 5G technology, specifically looking at the use of segment routing technology, and SRv6 in particular. We also discuss the use of a traditional Software-defined networking (SDN) architecture for SRv6 technology, which involves a centralized decision-making process for Segment Lists (SL), which are applied to provide services [3].

Next, we turn our attention to the platforms and programs used in our own network simulation. We provide short summaries of five different tools related to networking and network analysis that were used to create our simulation testbed. These tools include FD.io VPP, Docker, Trex, Network Configuration Protocol (NETCONF)/Ansible and Tshark/Tcpdump.

Finally, we present an overview of the topology we created for our simulation. The architecture of the mobile transport network is established according to the standard 5G transport network design, with a simplified diagram presented in Fig. 1. We tested this testbed by integrating real-enviroment data patterns to obtain results for follow-up research.

Overall, the research presented in this paper highlights the importance of combining the records from the real mobile network environment with network simulation in the development and optimization of mobile 5G transport networks. By exploring different tools and techniques, we hope to provide insights and outputs that could update on demand the routing strategies in the transport network. Moreover, in future research results from your simulations can be used to predict the network behaviour to optimize the transport network.

II. Background and Related Works

In this section, we describe research activities related to network simulations in connection with mobile 5G transport network and integrating real-environment data into the topology. The use in mobile networks is the reason why network simulation is specific and different from classical network simulation. The paper considers the future use of network simulation for further research and therefore considers the use of network segment routing technology, SRv6 specifically. Which seems to be the solution in IP networks for 5G network slicing technology.

One of the articles that considers the use of FD.io VPP is [1]. Authors contemplate that there are two main open source data plane implementations in case of SRv6 dataplane for software routers implementation: the Linux kernel and the FD.io VPP project. According to their research FD.io VPP is a platform that provides out-of-the-box switch/router functionality that can run on commodity CPUs. It offers better support for SRv6 Policy Headend and endpoint behaviors compared to the Linux kernel. VPP also has the capability to create and associate Binding Segment ID (SID) with an SR policy, while the Linux kernel lacks this support.

The article [2] suggests using a traditional SDN architecture for SRv6 technology. This involves a centralized decision-making process for Segment Lists, which are applied to provide services. The SDN controller uses a southbound Application programming interface (API) to communicate with SR-enabled devices and enforce the application of Segment Lists. To facilitate development and testing, the authors have created an Intent-based emulation system, which can be downloaded from their project page [3]. However, this emulated system does not meet the requirements for entering real 5G traffic.

In [4] authors focus exactly on 5G transport network, where they introduced SDN-enabled framework which collects network topology information in a centralized way and analyzes network information for network path computation. The experiment conducted aimed to evaluate the proposed framework in a simulated Chunghwa Telecom 5G backhaul network using the K-Shortest Path Algorithm. The network environment had more than 900 network nodes and 4000 unidirectional links between them, and the proposed framework was able to handle the large-scale network topology. Their simulation therefore had a different character and did not use classical network simulation or emulation tools.

Classical simulation and emulation network tools are considered in [5], [6] and [7]. Unfortunately, they lack full SRv6 support and the ability to create and implement new features in any way.

5G Mobile Data Patterns has been investigated in [8], authors use 5G trace dataset collected from a major Irish mobile operator. The dataset is composed of client-side cellular key performance indicators (KPIs) comprised of channel-related metrics, context-related metrics, cell-related metrics and throughput information. To supplement their real-time 5G production network dataset, they provide a 5G large scale multi-cell ns-3 simulation framework. However, the authors do not consider the use proposed in this paper, which is the use of mobile traffic patterns to find the delay on links in mobile transport network.

III. Utilized platforms and programs

This paper presents short summary of five different tools related to networking and network analysis. These tools were used to create a simulation testbed, which serves as the basis for all measurements and analyses. These tools are FD.io VPP, Docker, Trex, NETCONF/Ansible and Tshark/Tcpdump.

FD.io VPP provides a high-performance, packet-processing framework for networking applications. VPP is designed to run on commodity hardware and offers a range of features for routing, switching, and other network functions. VPP support for SRv6 Policy Headend and endpoint behavior. These behaviors are implemented in dedicated VPP graph nodes, and whenever an SRv6 segment is instantiated, a new IPv6 FIB entry is created for the segment address pointing to the corresponding VPP graph node. Furthermore, FD.io VPP supports the concept of SR policy. An SR policy is uniquely identified by its BindingSID address, which serves as a key to a particular SR policy. This allows for more efficient and flexible steering of traffic in the network [9].

Docker is a containerization platform that offers flexible and efficient solutions for deploying network services and applications. In the context of routing, Docker allows for the creation of lightweight and isolated containers that can run routing protocols and other network services. Docker's networking feature provides a range of options for configuring and managing network connectivity within and between Docker containers, enabling network administrators to easily deploy and manage network routing services in a scalable and reliable way [10].

Trex is an open-source software application that provides a platform for generating and analyzing network traffic. With its high-performance packet generation capabilities and support for a variety of network protocols. In recent years, Trex has been integrated with Docker to provide a more flexible and scalable platform for generating and analyzing network traffic. The integration of Trex with Docker provides a powerful and efficient platform for generating and analyzing network traffic in a flexible and scalable way [11].

NETCONF is an IETF standard protocol used for configuring and managing network devices. Ansible is a popular open-source software tool used for automating configuration management and orchestration tasks. In VPP, NETCONF can be used to configure and manage the routing and other network functions. By integrating NETCONF with Ansible, network administrators can automate the configuration and management of VPP instances. Ansible provides a declarative and idempotent way of defining the desired state of VPP instances and NETCONF makes it possible to configure and manage the network devices [12].

Tshark is a command-line network protocol analyzer that is part of the Wireshark network analysis software suite. It is a tool for capturing and analyzing network traffic in real-time or from saved capture files. Tshark supports a wide range of network protocols and can be used to generate statistics, filter traffic, and decode packets. Tshark's command-line interface provides flexibility and ease of use, making it an ideal choice for use in scripting and automation. Additionally, Tshark's support for a variety of capture file formats and export options makes it a versatile tool for sharing and analyzing network captures [14].

Tcpdump is an open-source command-line packet capture tool that is used for network traffic analysis. It allows network administrators to capture and examine network packets in real-time, providing a way to troubleshoot network issues, monitor network performance, and detect security threats. Tcpdump's lightweight design and low system resource requirements make it ideal for use in a wide range of environments, including cloud, IoT, and edge computing [13].

IV. Overview of Created Topology

The architecture design of the mobile transport network is established according to the standard 5G transport network design [15]. A simplified diagram of the architecture is presented in Fig. 1, which illustrates the Radio Access Architecture (RAN) on the left, followed by a topology comprising six routers. The middle four routers are connected to each other in a fully meshed network, while the 5G core is located on the right. The Cell Site in the simulation represents the Trex generator and the 5G core, serving as a transmitter and receiver.

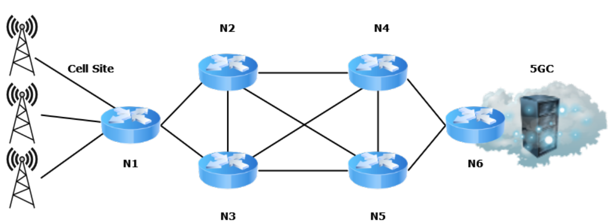


Fig. 1. Example of a simplified architecture diagram for a mobile transport network

The data were modified to reflect individual network slices, as defined by the 3rd Generation Partnership Project (3GPP) Mobile Transport Network Components Identification (MTNC-ID) [15]. SRv6 technology is assumed to be used by all routers in the topology, hence VPP of FD.io is utilized.

To enable centralized access and management of the topology, Docker software was employed, which provides tools to create, deploy, and run containers using simple commands or automation via APIs. Using Docker, an Ubuntu instance was started, on top of which FD.io VPP was initiated, while the Trex generator was launched by Docker. The FRR routing protocol was used to enable communication between Ubuntu nodes, which directly transmit data to the VPP instance. The VPP platform can be configured via the NETCONF protocol thanks to the Honeycomb agent, which allows NETCONF messages to be processed.

The data is encapsulated in SRv6, where it is assigned with SRv6 policy and a final destination based on the header information in the received packet. The packets generated by the Trex generator are modified in the header to simulate the creation of an MTNC-ID, which is carried in the SRv6 header. Whole created setup is presented in Fig. 2.

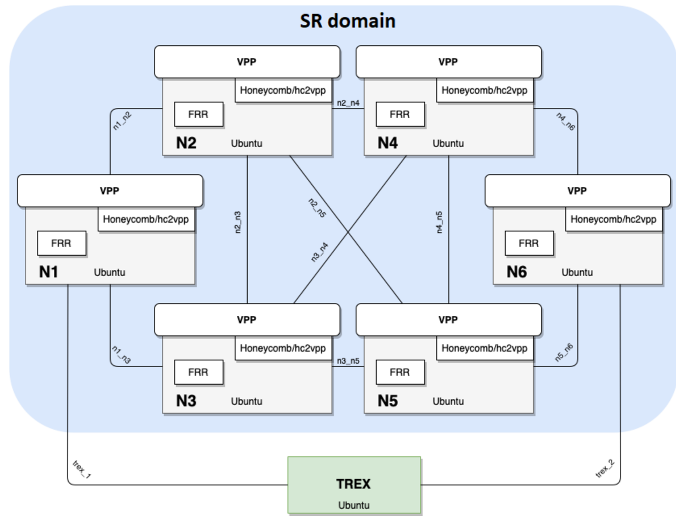


Fig. 2. Topology built on top of the Docker platform

All configuration is loaded by Ansible at startup. The actual configuration of the VPP nodes is then uploaded using playbooks again using Ansible. This makes it possible to freely change the configuration as needed.

Verification of the full functionality of the topology can be observed using the Trex console tui, where there is no packet drop and the transmitted data can be observed, in the Fig. 3.

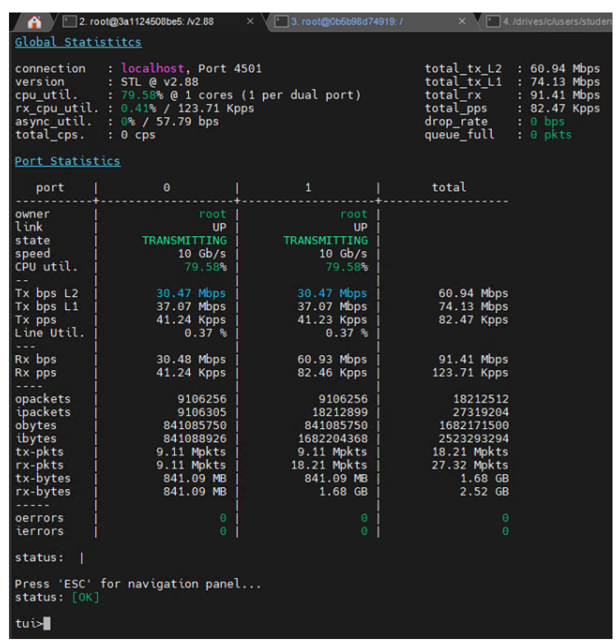


Fig. 3. Verification of topology functionality using Trex tui

The created environment is generic, any configuration can be set on it. We used it to integrate real-environment patterns from the data described in the next paragraph.

V. Description of the real data

The dataset comprises network quality records gathered from three distinct eNodeBs consisting of remote radio head (RRH)+ Baseband Unit (BBU), belonging to one cell sites. Over the span of 200 days, each eNodeB has been monitored and 1200 log records consisting of hourly observations of mobile data usage have been collected. The cell site serve a total of 53 000 users on average with 375 gigabytes of traffic volume per day. Table I summarizes key physical communication parameters of the cell sites.

TABLE I
Physical parameters of data

Parameters	Values
City size	300 km^2
Type of area	Residential
Population density	0.5 mil people
Cell type	Macro
Number of cell sites	1
Number of eNodeBs (RRH+BBU)	3
Each eNodeB	6 sectors - 12 cells
Observed time	7 months/every hour

The data is collected from one macro cell site, situated in a residential area in cell site shelters between residential buildings. The distance between this cell site and another one in that area is approximately 3 km. The precise locations of the base stations have been intentionally omitted to preserve the commercial benefits of the operator. The data has been recorded via an analytical platform which monitors the performance of cell sites in the network and provides other relevant information concerning eNodeBs such as cell site status, the number of drops, blocks, etc. Every hour, each eNodeB sends key performance indicators (KPIs) to this platform database in the form of raw counters. The collected data includes information about data volumes in the downlink and uplink directions, the number of connected UEs, and packet loss. Data are completely anonymous. The volume part plays key role in defining the traffic patterns, that are inserted through Trex into the created topology.

VI. Enhancing network traffic based on real data set in created topology

Insertion of data into the topology is achieved using the Trex real traffic generator. The packets themselves are thus generated using this program. Specifically we used the IMIX profile. IMIX profile is a stream of artificially created UDP packets of specific size. In this profile is it also possible to set up parameters such as packets per second (pps) and inter-stream gap (ISG).

The size of the generated packet has a fixed structure and is set to 160 bytes, so that it corresponds to the size of the packet for voice communication [16]. Thanks to this parameter outputs ready for the optimization for Ultra-Reliable Low Latency Communications (URLLC) slice is targeted. Pps was set to 10 and ISG 0.2 s.

Real data described in previous chapter were measured for 200 days. In case of our simulation we only use 7 days, one whole week, because of the time consumption of overall measurement. From this real data we used the patterns that we applied to the artificially created UDP packets in the Trex. Then we performed measurements, when the data with the patterns were continuously fed into the topology over 7 days.

A. *Transferring Empirical Data Patterns to Simulation Models*

The raw counters collected from the eNodeBs were in hourly interval. We were forced to break these hourly data volumes into smaller time periods to better simulate fluctuations in the network. So that we can better simulate the problems faced by transport networks such as loss rate, insufficient throughput or high delays. In mobile networks, traffic fluctuations can occur due to the movement of people.

The rate of generated packets changed every 5 min. However, the total amount of generated data was equal to the hourly volume measured in the real network. A script was created to generate this formula. The size of the UDP packet and the total data volume during one hour were taken into account. The resulting values for 5 min intervals were then fed into the Trex generator using the command `update -m x mbps`, again using the script.

B. *Measured data capture and processing*

While the data was being sent to the topology, it was necessary to capture behavior on the network. This was achieved by developed scripts. The capture time was every 15 min for 15 s. 15 min is the interval during which the KPIs are collected on the eNodeBs. So our measurement in this corresponds to the behavior in a real mobile network. 15 s was chosen to save the size of the final file and to facilitate post-processing. A measurement for 60 s was also tested, but did not yield significant differences and therefore 15s was retained.

During these 15 s, the time-stamped packets being transmitted were captured by TCPDump at the VPP. These captured packets were stored in pcap format. In order to further work with these pcap files, which were 96 per one day (24h x every 15 min for 15 s), the TShark program was used. A total of 672 records were measured throughout the week. Using TShark, the data was cleaned and only parameters related to the use for network slicing were selected. Delay for URLLC, total amount of data for Enhanced Mobile Broadband (eMBB) and packet loss for Massive Machine-Type Communications (mMTC). Now we focused primarily on delay outputs in this measurement.

The timestamp in the packets was taken into account at the moment the packet arrived at the VPP and at the moment it left the VPP after it was processed. This process was on all routers. By capturing packets with timestamps, we are able to determine the delay on the link between two routers. It always takes the packet arrival time on one router minus the packet departure time on the next router along the path. This logic is shown graphically in the Fig. 4.

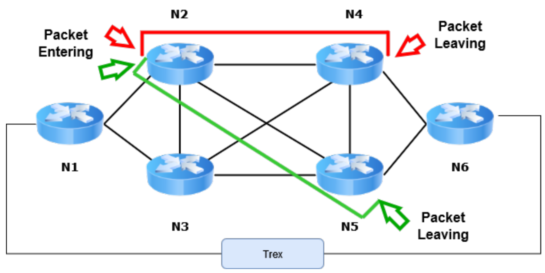


Fig. 4. Graphical representation of packet capture locations

Source IP	*Next hop SID*	*Size [B]*	*SID left [-]*	*Time*	*Epoch Time [s]*		
c000::223b:6467	fd44::100	248	3	Jan 25, 2023 12:30:28.353707500 UTC	1674649828		**Router N2**
d000::223b:6467	fd44::100	248	2	Jan 25, 2023 12:30:28.353710100 UTC	1674649828		
a000::280d:5a54	fd22::100	232	3	Jan 25, 2023 12:30:28.357127600 UTC	1674649828		
c000::280d:5a54	fd22::100	248	4	Jan 25, 2023 12:30:28.357137600 UTC	1674649828		
a000::280d:5a54	fd55::100	232	2	Jan 25, 2023 12:30:28.357233500 UTC	1674649828		
c000::280d:5a54	fd44::100	248	3	Jan 25, 2023 12:30:28.357246300 UTC	1674649828		
d000::280d:5a54	fd22::100	248	3	Jan 25, 2023 12:30:28.357358100 UTC	1674649828	enter N2	←
d000::280d:5a54	fd44::100	248	2	Jan 25, 2023 12:30:28.357648200 UTC	1674649828	leave N2	
c000::1c1e:14d7	fd22::100	248	4	Jan 25, 2023 12:30:28.361000700 UTC	1674649828		
c000::280d:5a54	fd44::100	248	3	Jan 25, 2023 12:30:28.357249200 UTC	1674649828		**Router N4**
b000::280d:5a54	fd44::100	232	2	Jan 25, 2023 12:30:28.357364700 UTC	1674649828		
e000::280d:5a54	fd44::100	248	2	Jan 25, 2023 12:30:28.357481700 UTC	1674649828		
c000::280d:5a54	fd55::100	248	2	Jan 25, 2023 12:30:28.357506800 UTC	1674649828		
b000::280d:5a54	fd66::106	232	1	Jan 25, 2023 12:30:28.357511800 UTC	1674649828		
e000::280d:5a54	fd66::100	248	1	Jan 25, 2023 12:30:28.357516400 UTC	1674649828		
d000::280d:5a54	fd44::100	248	2	Jan 25, 2023 12:30:28.357652500 UTC	1674649828	enter N4	
d000::280d:5a54	fd66::100	248	1	Jan 25, 2023 12:30:28.357729800 UTC	1674649828	leave N4	←
c000::1c1e:14d7	fd44::100	248	3	Jan 25, 2023 12:30:28.361061500 UTC	1674649828		

Fig. 5. Sample measured records for further research with capture logic

The collected data example with the logic and selected data information, can be seen also in Fig. 5.

These records are important for analysing network behaviour. Thanks to these records we are able to monitor the parameters in the network (in this case delay) and based on it change the routing in the network to create load balancing for network slice, that is being transmitted. As you can see in Fig. 5, we are able to calculate the delay from the captured packets at router N2 and N4. Similarly, we are able to evaluate the delay on all other links to determine the ideal routing for a given slice type. Since there is a large number of records, the cleaning and processing of data is a separate problem and it is dealt with in another article. Here we only present the method of measurement in created topology and data collection.

VII. Conclusion

In this text, we have explored the use of network simulation in the context of mobile 5G transport networks. Our analysis begins by describing related research activities, including efforts to integrate real-world environment data into network simulations.

In our analysis, we focused on the platforms and programs used in our own network simulation. We provided short summary of five different tools related to networking and network analysis that were used to create our testbed.

In this paper we introduce designed and created a mobile transport network topology based on the 5G standard, which contains a network of six routers using SRv6 technology and it is centrally configurable using Ansible.

In this article we have also presented real data from a macro cell site in a residential area, consisting of hourly observations of mobile data usage over the span of 200 days. This data was fed into the topology using the Trex, which was set up to simulate real-world traffic patterns over a period of 7 days. We created a script that reduced the hourly intervals to shorter periods of time to fluctuate the changes in the network. This resulted in 672 records of 15 s intervals with information about the network delay for a certain type of network slice. Which serve as output for further research.

We can create a new routing that is more optimized for the slice type based on our outputs. This logic could bring many advantages for mobile network operators that would consider applying network slicing technology in the transport part of the network. In addition, the follow-up research offers the possibility of predicting the overall network behaviour based on the integration of these real data and the use of machine learning. So we are able to estimate in advance how the network setup should look like.

References

[1] P. L. Ventre et al., "Segment Routing: A Comprehensive Survey of Research Activities, Standardization Efforts, and Implementation Results," in IEEE Communications Surveys & Tutorials, vol. 23, no. 1, pp. 182-221, Firstquarter 2021, doi: 10.1109/COMST.2020.3036826.

[2] P. Ventre et al., "SDN Architecture and Southbound APIs for IPv6 Segment Routing Enabled Wide Area Networks," IEEE Transactions on Network and Service Management, 2018..

[3] SRv6-SDN. [Online]. Available: https://netgroup.github.io/srv6-sdn/

[4] H. -C. Chi, W. -H. Chang, M. -C. Tseng, Y. -P. Yu and K. -H. Lai, "SDN-enabled Framework for Resilient Traffic Routing in 5G Networks," 2022 23rd APNOMS, Takamatsu, Japan, 2022, pp. 1-4, doi: 10.23919/APNOMS56106.2022.9919952.

[5] T. Tamanna and T. Fatema, "MPLS VPN over mGRE design and implementation for a service provider's network using GNS3 simulator," 2017 WCSP, WiSPNET, Chennai, India, 2017, pp. 2339-2342, doi: 10.1109/WiSPNET.2017.8300178.

[6] E. Zeydan, J. Mangues-Bafalluy and Y. Turk, "Intelligent Service Orchestration in Edge Cloud Networks," in IEEE Network, vol. 35, no. 6, pp. 126-132, November/December 2021, doi: 10.1109/MNET.101.2100214.

[7] A. A. Barakabitze, L. Sun, I. -H. Mkwawa and E. Ifeachor, "A Novel QoE-Centric SDN-Based Multipath Routing Approach for Multimedia Services over 5G Networks," 2018 IEEE ICC, Kansas City, MO, USA, 2018, pp. 1-7, doi: 10.1109/ICC.2018.8422617.

[8] Darijo Raca, Dylan Leahy, Cormac J. Sreenan, and Jason J. Quinlan. 2020. Beyond throughput, the next generation: a 5G dataset with channel and context metrics. In Proceedings of the 11th ACM MMSys'20. Association for Computing Machinery, New York, NY, USA, 303–308. https://doi.org/10.1145/3339825.3394938

[9] FD.io. VPP [Online] [https://wiki.fd.io/view/VPP].February 2023.

[10] Docker. Docker overview [Online] [https://docs.docker.com/get-started /overview/]. February 2023.

[11] Cisco.TRex[Online][https://trex-tgn.cisco.com/trex/doc]. February 2023.

[12] Enns, R., Björklund, M., Bierman, A. a Schönwälder, J. Network Configuration Protocol (NETCONF) [RFC 6241]. RFC Editor, Jun 2011. DOI:10.17487/RFC6241. Available from:https://www.rfc-editor.org/info/rfc6241.

[13] TCPDump. TCPDump home web site [Online] [https://www.tcpdump.org/]. February 2023.

[14] TShark. tshark.dev home [Online] [https://tshark.dev/]. February 2023.

[15] U. Chunduri, J. Kaippallimalil, S. Bhaskaran, J. Tantsura, and P. Muley, "Mobility aware Transport Network Slicing for 5G," IETF, Internet-Draft draft-ietf-dmm-tn-aware-mobility-05, Oct. 2022, work in Progress. [Online]. Available: https://datatracker.ietf.org/doc/draft-ietf-dmm-tn-aware-mobility/05/

[16] Ramalho, M., Ed., Jones, P., Harada, N., Perumal, M., and L. Miao, "RTP Payload Format for G.711.0", RFC 7655, DOI 10.17487/RFC7655, November 2015, ¡https://www.rfc-editor.org/info/rfc7655¿.

Enhancing Security Monitoring with AI-Enabled Log Collection and NLP Modules on a Unified Open Source Platform

1st Yehor Safonov
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic,
0000-0002-3549-2178
yehor.safonov@vutbr.cz

2nd Michal Zernovic
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
xzerno00@vutbr.cz

***Abstract*—The number of computer attacks continues to increase daily, posing significant challenges to modern security administrators to provide security in their organizations. With the rise of sophisticated cyber threats, it is becoming increasingly difficult to detect and prevent attacks using traditional security measures. As a result, security monitoring solutions such as *Security Information and Event Management* (SIEM) have become a critical component of modern security infrastructures. However, these solutions still face limitations, and administrators are constantly seeking ways to enhance their capabilities to effectively protect their cyber units. This paper explores how advanced deep learning techniques can help boost security monitoring capabilities by utilizing them throughout all stages of log processing. The presented platform has the potential to fundamentally transform and bring about a significant change in the field of security monitoring with advanced AI capabilities. The study includes a detailed comparison of modern log collection platforms, with the goal of determining the most effective approach. The key benefits of the proposed solution are its scalability and multipurpose nature. The platform integrates an open source solution and allows the organization to connect any event log sources or the entire SIEM solution, normalize and filter data, and use this data to train and deploy different AI models to perform different security monitoring tasks more efficiently.**

***Index Terms*—Artificial intelligence, deep learning, Fluentd, log collection, log processing, Logstash, security monitoring, SIEM.**

I. Introduction

In today's world, it is impossible to imagine any modern field without a computer infrastructure [1]. Given its ubiquity, ensuring uninterrupted and secure operation of these systems has become a top priority [2], [3]. Various hardware and software solutions are utilised inside cyberspaces to achieve their smooth functioning. When it comes to computer infrastructure, there are four main categories of elements that could be found: *Networking Devices* (such as routers, switches, and firewalls), *Operating Systems* (including Unix-based and Windows systems), *Security Applications* (such as Next Generation Firewalls (NGFW), Vulnerability Management Systems (VMS), Endpoint Detection and Response (EDR), Data Los Protection (DLP), Privileged Access Management (PAM), etc.) and *Other Applications* (such as database systems, customer applications, cloud applications, etc.) [3]. All of the above elements help to ensure the smooth running of the organisation and enable all organisational processes to be fulfilled.

These devices, along with end devices and applications, generate records of their traffic and usage, known as logs [2], [4]. They are crucial for evaluating the network's status, security, and events that occur on it. They can provide insight into potential attacks on the infrastructure or can be used to analyse previous incidents and take appropriate action [1]. The log analysis process is essential for discovering relevant information from these records. From a security perspective, all log records must be centrally analysed, normalised, and stored for the long term to ensure effective security monitoring [3]. The main challenge of this issue is the lack of standardisation on the LES (*Log Event Source*) level due to the existence of different log formats and protocols for their transmission, see Fig. 1. This complicates their central processing.

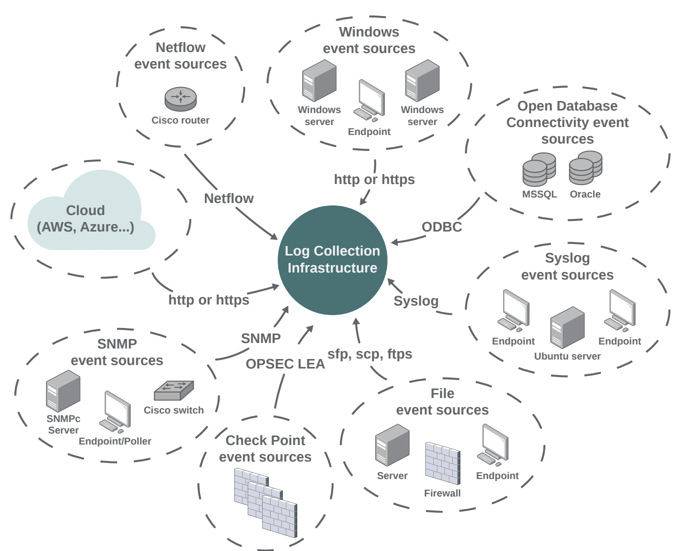


Fig. 1. Schema of centralized log collection in computing infrastructures.

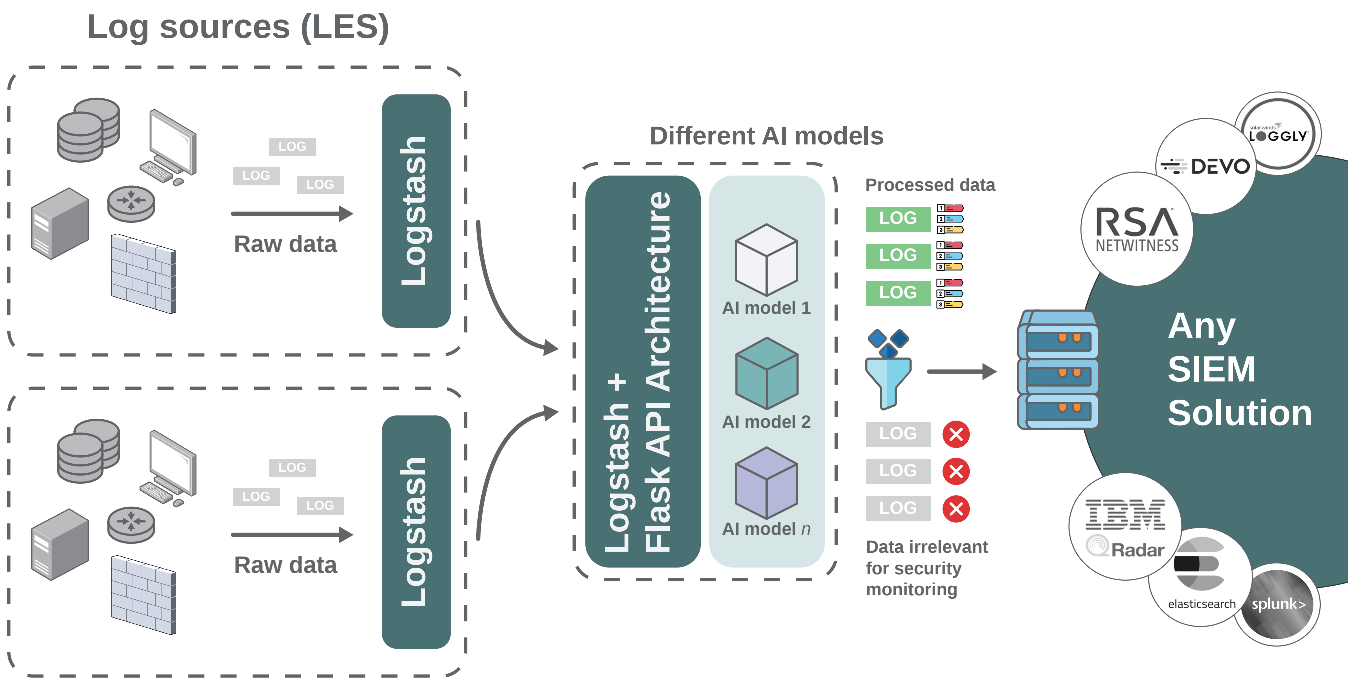


Fig. 2. A way of integration purposed platform into log processing to cooperate with SIEM systems, while also demonstrating its scalability.

In general, there are two distinct methods for providing security monitoring. Simple analysis techniques, such as manual commands to retrieve records, may be suitable for smaller infrastructures, but are not effective for larger systems [4]. In this day and age, advanced analysis techniques are offered within SIEM systems, which also allow responding to security incidents, creating rules, and notifying and preventing attacks. SIEM offers capabilities such as advanced log analysis, log storage, and other advanced features [2].

One way to enrich log processing is to embed a neural network into it. Although current proprietary SIEM solutions use neural networks to solve correlation tasks, the presented approach performs a wide range of security tasks and allows SIEM systems to be more efficient.

II. On current systems for security monitoring

SIEM is a platform or a system that is responsible for collecting security-relevant data in a centralised environment to detect threats and incidents. This system consists of two concepts: SIM is used to collect security-relevant data and generate reports. SEM provides security incident analysis, incident correlation, and real-time alerting. Together, these two categories constitute a SIEM. [2]

Before the records can be analyzed, a series of steps encompassing various SIEM functions must be completed within the SIEM tools: [3]

Log collection – can be done in different ways. Some log sources are based on a push technique (e.g. Syslog), where the user on the end device configures the destination IP address and port. The user then configures a listener on the desired device and receives the logs. Some sources in turn work on a pull technique (e.g. Windows event log), where it is necessary to connect to the machine and download the logs. [3], [5], [6]

Normalization – all log forms require specific connectors. They convert different log file formats from different devices, versions to a common SIEM understandable format. [1]

Aggregation – a process of removing duplicates. Logs are aggregated according to similar characteristics, thus opening up the possibility to discard unnecessary data. [3]

Correlation and analysis – a great advantage of these solutions are the correlation rules. A constant stream of logs from different connected devices should flow into the SIEM. Correlation rules tell the SIEM which sequence of events indicates a potential incident or vulnerability. [3]

To illustrate the concept, the following example is useful: alert administrators if the same IP address attempts to log on to a particular device on the network. It fails five times with different login credentials and then successfully connects to a device on the network. All events occurred within a fifteen-minute time window. This sequence of events may be an indication that a potential attacker has gained authentication into the network and could signal an escalation of privilege, so a potential security incident occurred. [7]

When setting up a SIEM system, the rules in question must be reprogrammed as needed. They are inserted manually, and therefore, it is necessary to know which anomalies do not make sense to report and which should trigger an alarm. [4]

Notification and response – In the event of a security breach, SIEM alerts the personnel responsible for network protection in real time. It also selects an adequate response for the optimal protection of assets. [3]

It can be observed that the SIEM is a complex system with a lot of functionality. The entire operation of receiving and processing logs can be enriched by integrating a neural network to perform various security tasks.

As seen in Fig. 2, the way it would work is as follows: Different locations contain a certain infrastructure that produces many logs, which can be considered for this purpose as "raw data". The infrastructure would forward these data to a Logstash instance, Logstash being the software to collect these logs. Log collectors allow for scalability, which is useful for not only distributing the load but also creating logical units by separating locations or infrastructures. Logstash then sends these data to a primary Logstash instance, which directly communicates with the platform. The platform would then be responsible for routing logs into the AI model, which performs various advanced security tasks. This makes it much easier to deploy this type of infrastructure, and the AI networks add many advantages to the logging process.

III. The advantages of AI in log processing

Technological advancements constantly push the boundaries of what is possible, with new inventions and innovations emerging on a daily basis. Throughout history, transformative breakthroughs such as the steam engine, electricity, computers, and the Internet have revolutionised society and propelled human progress forward. These advances have fundamentally transformed existing processes and pushed society to new heights on the evolutionary spiral [8].

The fourth industrial revolution has been driven primarily by advances in artificial intelligence and machine learning, which enable intelligent automation and data-driven decision-making [8], [9]. Artificial intelligence has been used in different fields to improve existing processes and uncover previously unknown relationships between data sets. Artificial intelligence techniques are proving to be highly promising for security monitoring, especially in log processing [9].

Log records which are processed by SIEM solution are represented as unstructured text data, as there are a multitude of log event sources and formats. Since log records often contain a large amount of semantic information that cannot be captured by traditional statistical methods, it is possible to use *natural language processing* (NLP) algorithms to tackle various security tasks [8]. There are a variety of options in the context of processing log data effectively and provide security monitoring. These include, but are not limited to, log correlation to identify patterns, log anonymisation to protect sensitive information, log parsing to extract useful data, and log filtering to remove irrelevant data with the motivation to optimise SIEM licence usage. In order to mitigate this problem, deep artificial networks can be used [8], [10]. However, it is worth noting that the latter may have a constraint on the maximum token size for the input sequence, despite achieving the state-of-the-art results for various NLP tasks such as Question Answering, Classification, Categorisation, and more.

Moreover, the popularity of AI algorithms is further bolstered by the limitations of modern SIEM solutions. Due to the limited ability of these systems to respond automatically to cybersecurity incidents, without human intervention. Correlation rules must be mainly written manually, making it impractical to address every possible security incident. In addition, there may be a lack of the data needed to process and detect all security incidents effectively. The usage of archived data during incident investigation can also pose a challenge. [11]

One of the main challenges in using AI techniques for log processing is the requirement for fast data processing. This is because modern SIEM solutions operate on real-time log data and must evaluate potential attack vectors as quickly as possible. It is important to know that modern SIEM solutions typically have a data flow rate of more then 5000 EPS[1], which is based on the number of connected LESes. [12]

IV. Comparison of log collection approaches

Multiple approaches can be taken when creating a platform which supports the integration of AI models into log processing. These approaches vary in difficulty and effectiveness.

One of them is to create a log collector integrated within the platform. A simple collector can be made using a programming language like Python and its various libraries, which allow the manipulation of logs. Unfortunately, making a reliable and safe collector that is capable of pulling logs from many different sources would be a very lengthy and difficult process. There are already log collectors that can be used for this purpose and connect into the platform.

This section contains a comparison between three popular choices when it comes to log collecting. Each of them has its own advantages and fits differently to different architectures.

A. *NXLog*

This collector is supported on many different operating systems, which is its biggest advantage when compared to other collectors. It is the only application in this section that also has a paid version [13]. A configuration file is what is used to control the flow of logs into the application. The file consists of directives that are similar to XML tags. [14]

The process of creating a file consists of defining the constants, setting the global directives, input, output and route configurations [14]. This is a lengthier process when compared to other mentioned collectors.

B. *Fluentd*

An open source collector, which has many plugins to support a wider range of protocols and offers different options to work with LESes. There are 9 types of plugins: Input, Parser, Filter, Output, Formatter, Storage, Service Discovery, Buffer, and Metrics. [15]

The heart of this software is a configuration file, where the main parts are the input and output plugins. Plugins are defined in the file, and within them there are tags that specify information such as ports, protocol types, and other configuration fields. [15]

[1]EPS (*Events per second*) refers to the rate at which events (such as log entries) are generated and processed by the system. The EPS rate is a measure of how much log data a SIEM solution can handle and process in real-time.

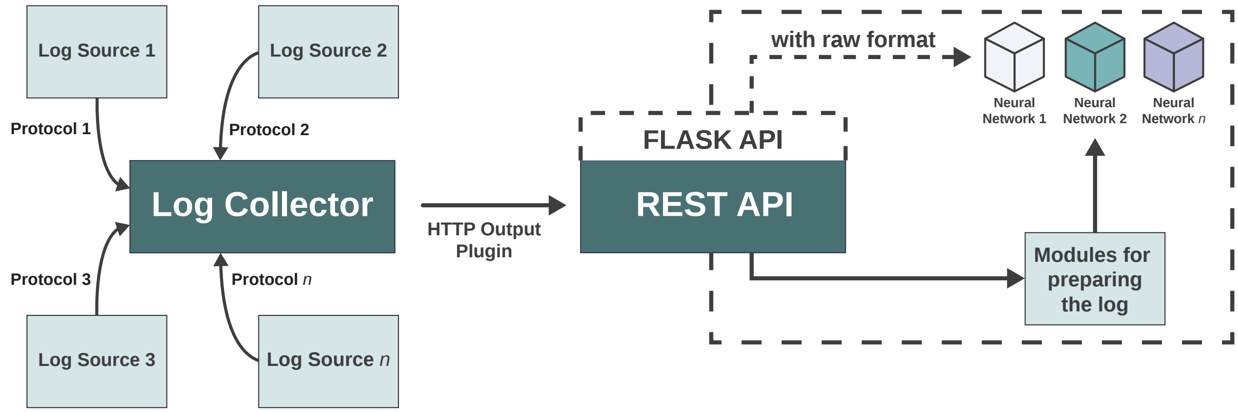


Fig. 3. Simplified view of the architecture of the proposed platform.

C. Logstash

Logstash is an open source application which is part of the Elastic stack. Such a program allows for the collection of logs from different sources using inputs in the program configuration file, which is called a logstash pipeline. Pipeline can contain multiple fields, such as input, filter, and output, but only needs input and output to function. [16]

The input field contains the type of protocol used for receiving logs. It usually specifies the port on which the transfer will take place. [17]

When it comes to output, multiple formats are supported. Logstash supports many different log protocols, for example, syslog, file, redis, exec, jdbc, http and many others. [18]

Logstash is a good choice for this infrastructure, considering it is open source, has options for scalability (instances can forward data to other Logstash instances) and the configuration file is quick to set up. It also has a large community, which opens up options for support. It also makes getting certain logs easier by supporting an application called Beats. It is a part of the Elastic suite of software, and makes particular systems able to send logs automatically, without having to pull them manually. [19]

D. Final comparison of selected log collectors

There are a few metrics for the comparison of log collectors. **Supported platforms** – Fluentd and Logstash are supported on devices with Windows and Linux operating systems. NXLog is supported on both platforms and also on macOS and many others, such as Oracle Solaris. [13], [19]

Event routing – Fluentd relies on tags for routing. Logstash, on the other hand, uses logical expressions and elements. NXLog uses blocks similar to XML tags. At its core, the syntax and main idea of the configuration are similar across the three applications. [19]

Support for extensions – while with Logstash the effort is to keep all plug-ins centralised under a single GitHub repository, Fluentd offers only 11 input plugins in its official repository; all the others are created outside of it and, therefore, have decentralised access. NXLog offers plug-ins, extensions, and, by default, modules; a list of which is available in the corresponding documentation. Table I lists some of the standard plugins supported in the compared applications. The number of supported plugins can be increased by using the repositories and extensions mentioned above. [19]

Transport and Reliability – in terms of data retention after reboot and a transfer of data, Logstash requires Redis, which serves as an external queue for data preservation. Fluentd by default includes options to configure the system for caching. NXLog directly supports flow control and a log queue that stores data on disk in case of failure. [19], [20]

Performance – Fluentd uses only 40 MB of RAM compared to Logstash, which uses 120 MB. The recommended size of RAM for NXLog is 250 MB. This difference can be negligible in a broader network. Logstash solves the load by installing Beats, which are often used on endpoints to send logs to a central Logstash ELK station. [19]

TABLE I
A COMPARISON OF COMMON PROTOCOLS USED IN LOG COLLECTORS.

Fluentd	Logstash	NXLog
`Tail (File)`	`File`	`File`
`Syslog`	`Syslog`	`Syslog`
`UDP`	`UDP`	`UDP`
`TCP`	`TCP`	`TCP`
`HTTP`	`HTTP`	`HTTP`
`Windows Event Log`	`Beats`	`Windows Event Forwarding`
`Unix`	`Unix`	`Unix Domain Socket`
`Exec`	`Exec`	`Exec`
`Sample`	`Generator`	`Testgen`
`Monitor agent`	–	–
`Forward`	–	`Internal`
–	`Redis`	`Redis`
–	`Kafka`	`Kafka`
–	`Github`	–
–	`JDBC (DB)`	`ODBC (DB)`
–	–	`macOS Unified Logging System`
...	...	...

V. Results and discussion

The aim of the project is to use an open-source solution for log collection that supports scalability and a wide array of log sources. The inclusion of a reliable and free-to-use log collector is crucial to creating manageable pipelines and zones for data to flow through. The nature of logs and different protocols makes it necessary to use software that is capable of collecting and forwarding the data. After analysis, it was determined that Logstash is an application that meets the project requirements, considering its popularity, support, log types, and additional software. The next step is to create an API that can connect different AI models for different purposes. It is made to work in two modes, learning and real deployment. AI models are only as good as their input data, which means that collecting data for learning is crucial. Fig. 3 shows the design of the application. The log sources are configured to send logs to the Log Collector, which forwards the logs to the application programming interface via HTTP. It consists of modules to modify the format of the logs so that they are suitable for usage by the neural network. If the API receives the logs in a raw format, they can be directly pushed to the neural network without the need for processing modules to interfere with them. As shown in Fig. 2, the system is scalable and multiple Logstash instances can be connected to each other. In the case of real-time log processing, it is important to ensure that AI models can process data as quickly as possible and can handle a large number of EPS. The platform is not yet designed to be fully deployed into the cloud due to high data protection requirements.

VI. Conclusion and future work

The main goal of this paper was to analyse and enhance existing security monitoring processes using deep learning techniques. For example, modern deep neural networks like XLNet, BERT, and GPT2 can be transfer-learned and utilised to machinally understand the context of incoming log records and perform various NLP tasks. Discussed AI techniques revolutionised the field of deep learning by enabling parallel processing of input text data and avoiding recursion [21]. The main challenge of the modern security monitoring domain lies in the difficulty of accessing raw log source data and the lack of a platform that can seamlessly integrate normalised data with modern SIEM and Log Manager solutions. One of the significant contributions of this study is the creation of an open-source parallel platform based on Logstash that can seamlessly connect all modern log sources, normalise their outputs, and enable efficient training and deployment of any modern deep neural networks within the organisation's infrastructure. The Logstash solution was chosen after analysis of current log collection platforms, presented in section IV. Moreover, the presented solution can be successfully used for scientific purposes in the security monitoring domain. In the future, it is planned to optimise the discussed platform, deploy a prototype in real world infrastructures, and address the challenge of cloud deployment.

References

[1] BHATT, Sandeep, Pratyusa MANADHATA a Loai ZOMLOT. *The Operational Role of Security Information and Event Management Systems* [online]. 12. 2014 [cit. 26. 02. 23]. ISSN 1558-4046. Available at: https://dx.doi.org/10.1109/MSP.2014.103.

[2] Security Information and Event Management (SIEM). In: VIELBERTH, Manfred. *Encyclopedia of Cryptography, Security and Privacy* [online]. Springer: Springer, 2021, s. 1-3 [cit. 25. 02. 23]. ISBN 978-3-642-27739-9. Available at: https://doi.org/10.1007/978-3-642-27739-9_1681-1.

[3] HRISTOV, Marian, Maria NENOVA, Georgi ILIEV a Dimiter AVRESKY. Integration of Splunk Enterprise SIEM for DDoS Attack Detection in IoT. In: *2021 IEEE 20th International Symposium on Network Computing and Applications (NCA)*. Boston: IEEE, 2021, s. 1-5. ISBN 978-1-6654-9550-9. ISSN 2643-7929. Available at: http://dx.doi.org/10.1109/NCA53618.2021.9685977.

[4] MARTINASEK, Zdenek. Logging problematics, IDS and IPS systems [presentation]. ICT Security 2. Brno University of Technology, 2022.

[5] Installation and configuration for Windows Remote Management. *Microsoft* [online]. Redmond: Microsoft Corporation, © 2022, 10 september 2021 [cit. 26. 02. 23]. Available at: https://learn.microsoft.com/en-us/windows/win32/winrm/installation-and-configuration-for-windows-remote-management.

[6] Configuring rsyslog on a Logging Server. *Red Hat Customer Portal* [online]. Raleigh: Red Hat, © 2022 [cit. 26. 02. 23]. Available at: https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/6/html/deployment_guide.

[7] CRAWLEY, Kim. *How SIEM Correlation Rules Work*. AT&T Cybersecurity [online]. Dallas: AT&T, © 2022, 20 February 2018 [cit. 27. 02. 23]. Available at: https://cybersecurity.att.com/blogs/security-essentials/how-siem-correlation-rules-work.

[8] CHOLLET, Francois. *Deep learning with Python*. Shelter Island, New York: Manning Publications Co. [2018]. ISBN 1617294438.

[9] MINSKY, Martin. *Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. 2006. ISBN 0743276639.

[10] RUDD, Ethan a Ahmed ABDALLAH. *Training Transformers for Information Security Tasks: A Case Study on Malicious URL Prediction*. [online]. [cit. 11. 03. 23]. Available at: https://arxiv.org/pdf/2011.03040.pdf.

[11] GONZALES-GRANADILLO, Gustavo, Susana GONZALES-ZARZOSA a Rodrigo DIAZ. *Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures*. [online]. Madrid: Cybersecurity Unit, Atos Research & Innovation, 2021 [cit. 11. 03. 23]. ISSN 1424-8220. Available at: https://www.mdpi.com/1424-8220/21/14/4759.

[12] CONSTANTINE, Conrad. *What to log in a SIEM: SIEM and security logging best practices explained*. [online]. [cit. 10. 03. 23]. Available at: https://cybersecurity.att.com/blogs/security-essentials/what-kind-of-logs-for-effective-siem-implementation.

[13] OS Support. *NXLog* [online]. Newark: NXLog, © 2022 [cit. 02. 03. 23]. Available at: https://docs.nxlog.co/userguide/os/index.html.

[14] Configuration overview. *NXLog* [online]. Newark: NXLog, © 2022 [cit. 02. 03. 23]. Available at: https://docs.nxlog.co/userguide/configure/overview.html.

[15] Overview. *Fluentd* [online]. Fluentd Project, © 2022 [cit. 03. 03. 23]. Available at: https://docs.fluentd.org/quickstart.

[16] Support Matrix. *Elastic* [online]. Mountain View: Elasticsearch B.V., © 2022 [cit. 04. 03. 23]. Available at: https://www.elastic.co/support/matrix.

[17] Creating a Logstash pipeline. *Elastic* [online]. Mountain View: Elasticsearch B.V., © 2022 [cit. 04. 03. 23]. Available at: https://www.elastic.co/guide/en/logstash/current/configuration.html.

[18] Input plugins. *Elastic* [online]. Mountain View: Elasticsearch B.V., © 2022 [cit. 04. 03. 23]. Available at: https://www.elastic.co/guide/en/logstash/current/input-plugins.html.

[19] PERI, Noni. Fluentd vs Logstash: A Comparison of Log Collectors. *Logz.io* [online]. Tel Aviv: Logshero, © 2015-2022 [cit. 07. 03. 23]. Available at: https://logz.io/blog/fluentd-logstash/.

[20] Buffering and flow control. *NXLog* [online]. Newark: NXLog, © 2022 [cit. 02. 03. 23]. Available at: https://docs.nxlog.co/userguide/intro/buffering-and-flow-control.html.

[21] VASWANI, Ashish, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N. GOMEZ, Lukasz KAISER a Illia POLOSUKHIN. *Attention Is All You Need* [online]. [cit. 3. 3. 2023]. Available at: https://arxiv.org/pdf/1706.03762.pdf.

Gaussian Process Regression under Location Uncertainty using Monte Carlo Approximation

Martin Ptáček
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
xptace18@vut.cz
Brno Ph.D. Talent Scholarship Holder
Funded by the Brno City Municipality

***Abstract*—Gaussian Process Regression (GPR) is a common statistical framework for spatial function estimation. While its flexibility and availability of closed-form estimation solution after training are its advantages, it suffers on applicability constraints in scenarios with uncertain training positions. This paper presents the derivation of the exact GPR operating on uncertain training positions along with approximation of the resulting terms using Monte Carlo (MC) sampling. This method is then implemented in a simulation environment and shown to improve the estimation quality over the standard GPR approach with uncertain training positions.**

***Index Terms*—Spatial function estimation, GPR, Uncertain training positions, Probabilistic inference, Monte Carlo approximation**

I. Introduction

Spatial function estimation is a problem common to numerous scientific and practical fields including geostatistics, astronomy, meteorology, mining but also telecommunications and other machine learning applications. In the field of telecommunications the application can be found in estimating Channel state information maps to aid the development of mobile networks and optimization resource allocation [1] [2].

GPR is a method for spatial function estimation that recently gained popularity because of its success as a general machine learning method and its interpretation as a neural network with infinite basis functions [3]. The advantages of GPR are its flexibility and the availability of a closed form expression for the estimates after the training phase. Further significant advantages are the ability of GPR to indicate the level of belief (uncertainty) for the provided estimates and its performance on limited data sets.

GPR relies on precise knowledge of the training positions. In many practical scenarios, this may not be available and the performance of the GPR framework is prone to degradation. This motivates the search for approaches capable of overcoming this issue. This paper is focused on presenting the derivation of MC-GPR method, presenting the simulation setup, demonstrating the performance gains offered by this approach and the increased computation complexity costs.

II. Spatial Function Estimation

A spatial function in the chosen scenario may be imagined as a distribution of a scalar parameter over a continuous space. The space considered here shall be $D = 2$ dimensional, representing for example the Earth's surface. The task is to estimate the value of the spatial function at a given *test position* given the observations done by sensors placed at some given *training positions*. An example of such situation is depicted in Figure 1.

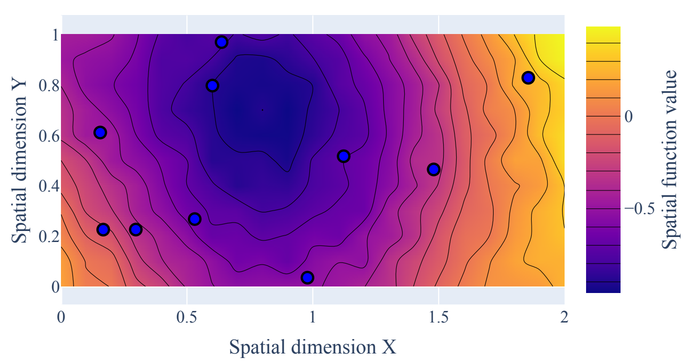


Fig. 1. Realization of a spatial function (represented by contour levels) and training positions (indicated by blue dots).

A popular framework for solving the Spatial function estimation problem is the Gaussian Process Regression.

III. Gaussian Process Regression Basics

A GP is a collection of random variables (RVs), any finite number of which have a joint Gaussian distribution.

The GP will be denoted as $f(\boldsymbol{x})$, where $\boldsymbol{x} \in \mathbb{R}^D$ is an argument specifying the position within the GP, in this case considered as a general position. The complete GP is then defined as

$$p(f(\boldsymbol{x})) = \mathcal{GP}(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \ , \tag{1}$$

The author would like to thank Ao.Univ.Prof. Franz Hlawatsch (TU Wien) for many helpful discussions on Gaussian Process Regression and early research guidance as well as to doc. RNDr. Jitka Poměnková, Ph.D. (BUT) for her comments.

Research described in this paper was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-20-6361.

DOI 10.13164/eeict.2023.222

where $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^D$ are generic position variables. Further, $m(\boldsymbol{x}) \triangleq \mathrm{E}\{f(\boldsymbol{x})\}$ is the *mean function* and $k(\boldsymbol{x}, \boldsymbol{x}') \triangleq \mathrm{cov}\{f(\boldsymbol{x}), f(\boldsymbol{x}')\}$ is the *covariance function* of the GP. $p(\cdot)$ here denotes *probability density function* but is further used also as *probability mass function* depending on the context.

A. Observation Model

We shall consider that the GP is observed at I known *training positions* $\boldsymbol{x}^{(i)}$ with $i = 1, 2, \ldots, I$ stacked to a vector according to

$$\boldsymbol{x}_\mathrm{t} \triangleq \mathrm{col}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(I)}) \in \mathbb{R}^{DI} \ . \tag{2}$$

The GP RVs $f(\boldsymbol{x}^{(i)})$ at the individual training positions $\boldsymbol{x}^{(i)}$ shall also be arranged into a vector according to

$$\boldsymbol{f} \triangleq (f(\boldsymbol{x}^{(1)}) \quad f(\boldsymbol{x}^{(2)}) \quad \ldots \quad f(\boldsymbol{x}^{(I)}))^\mathrm{T} \in \mathbb{R}^I \ . \tag{3}$$

It follows from (1) that $\boldsymbol{f}$ is a random vector with multivariate Gaussian distribution

$$p(\boldsymbol{f}) = \mathcal{N}(\boldsymbol{f};\, \boldsymbol{m},\, \boldsymbol{K}) \ , \tag{4}$$

where $\boldsymbol{m} = \mathrm{E}\{\boldsymbol{f}\}$ denotes the vector of the mean values and $\boldsymbol{K} = \mathrm{cov}\{\boldsymbol{f}\}$ the covariance matrix.

In the considered problem setup we assume not having direct access to the GP RVs $\boldsymbol{f}$ at the training positions. Instead, we are provided with vector $\boldsymbol{y}$ of observations of $\boldsymbol{f}$, i.e.,

$$\boldsymbol{y} \triangleq (y^{(1)} \quad y^{(2)} \quad \ldots \quad y^{(I)})^\mathrm{T} \in \mathbb{R}^I \ , \tag{5}$$

which is a noisy version of $\boldsymbol{f}$ according to

$$\boldsymbol{y} = \boldsymbol{f} + \boldsymbol{\epsilon} \ , \tag{6}$$

where $\boldsymbol{\epsilon} \triangleq (\epsilon^{(1)} \quad \epsilon^{(2)} \quad \ldots \quad \epsilon^{(I)})^\mathrm{T} \in \mathbb{R}^I$ is the vector of observation errors (measurement noise). $\boldsymbol{\epsilon}$ is assumed to be zero-mean isotropic Gaussian according to

$$p(\boldsymbol{\epsilon}) = \mathcal{N}(\boldsymbol{\epsilon};\, \boldsymbol{0},\, \sigma_\epsilon^2 \mathbf{I}_I) \ , \tag{7}$$

where σ_ϵ^2 is the single observation error variance and $\mathbf{I}_I$ is an identity matrix of dimension I.

It is assumed that $\boldsymbol{\epsilon}$ is independent of $\boldsymbol{f}$. From (6) and our statistical assumptions it follows that $\boldsymbol{y}$ is a Gaussian random vector distributed according to

$$p(\boldsymbol{y}) = \mathcal{N}(\boldsymbol{y};\, \boldsymbol{m},\, \boldsymbol{Q}) \tag{8}$$

with a covariance matrix $\boldsymbol{Q}$ of the random vector $\boldsymbol{y}$ given as

$$\boldsymbol{Q} = \mathrm{cov}\{\boldsymbol{f}\} + \mathrm{cov}\{\boldsymbol{\epsilon}\} = \boldsymbol{K} + \sigma_\epsilon^2 \mathbf{I}_I \ , \tag{9}$$

where we used the independence of random vectors $\boldsymbol{f}$ and $\boldsymbol{\epsilon}$.

B. Regression Formulation

The aim of the GPR is to estimate the GP RV at a single known *test position* $\boldsymbol{x}^{(*)} \in \mathbb{R}^D$, i.e., $f(\boldsymbol{x}^{(*)})$, which will be denoted shortly as f_*. To utilize the observations of the GP $\boldsymbol{y}$ at the training positions $\boldsymbol{x}_\mathrm{t}$ in order to estimate $f(\boldsymbol{x}^{(*)})$, the GPR framework expresses the posterior distribution (predictive distribution) of the RV to be estimated given the observed RVs according to the Bayes' Theorem as

$$p(f_*|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|f_*)\, p(f_*)}{p(\boldsymbol{y})} = \frac{p(f_*, \boldsymbol{y})}{p(\boldsymbol{y})} \ . \tag{10}$$

The likelihood term $p(\boldsymbol{y})$ is recognized as a normalizing constant. The posterior pdf can therefore be determined using the joint pdf $p(f_*, \boldsymbol{y})$ only, i.e., $p(f_*|\boldsymbol{y}) \propto p(f_*, \boldsymbol{y})$.

According to the definition of the GP, the joint distribution is Gaussian and given by

$$p(f_*, \boldsymbol{y}) = \mathcal{N}\left(\begin{pmatrix} f_* \\ \boldsymbol{y} \end{pmatrix};\, \begin{pmatrix} m_* \\ \boldsymbol{m} \end{pmatrix},\, \begin{pmatrix} k_* & \boldsymbol{c}^\mathrm{T} \\ \boldsymbol{c} & \boldsymbol{Q} \end{pmatrix} \right) \ , \tag{11}$$

where m_* and k_* are the prior mean and variance of f_*, i.e. the GP value at the test position, respectively. Vector $\boldsymbol{c} \triangleq \mathrm{cov}\{\boldsymbol{y},\, f_*\}$ is the cross-covariance of the vector of spatial function observations $\boldsymbol{y} = \boldsymbol{f} + \boldsymbol{\epsilon}$ and the GP random variable f_* at the test position $\boldsymbol{x}^{(*)}$. We obtain

$$\boldsymbol{c} = \mathrm{cov}\{\boldsymbol{f} + \boldsymbol{\epsilon},\, f_*\} = \mathrm{cov}\{\boldsymbol{f},\, f_*\} + \mathrm{cov}\{\boldsymbol{\epsilon},\, f_*\} \in \mathbb{R}^I \ . \tag{12}$$

Since $\boldsymbol{\epsilon}$ and f_* are independent, $\mathrm{cov}\{\boldsymbol{\epsilon},\, f_*\} = \boldsymbol{0}$ and thus we get $\boldsymbol{c} = \mathrm{cov}\{\boldsymbol{f},\, f_*\}$, which can be evaluated using the covariance function k as

$$\boldsymbol{c} = (k(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(*)}) \quad k(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(*)}) \quad \ldots \quad k(\boldsymbol{x}^{(I)}, \boldsymbol{x}^{(*)}))^\mathrm{T} \ . \tag{13}$$

Since the posterior distribution $p(f_*|\boldsymbol{y})$ is Gaussian, it is fully specified using the posterior mean $\mu_{f_*|y}$ and variance $\sigma_{f_*|y}^2$. These parameters can be expressed using the *completing the square* approach that expresses the posterior distribution according to the joint distribution (11) in a single exponential form, i.e., as a new Gaussian distribution

$$p(f_*|\boldsymbol{y}) = \mathcal{N}(f_*;\, \mu_{f_*|y},\, \sigma_{f_*|y}^2) \ , \tag{14}$$

where the posterior mean value is expressed as

$$\mu_{f_*|y} = m_* + \boldsymbol{c}^\mathrm{T} \boldsymbol{Q}^{-1} (\boldsymbol{y} - \boldsymbol{m}) \ . \tag{15}$$

This is the *Minimum Mean-squared Error* (MMSE) estimate of f_* given $\boldsymbol{y}$. The posterior variance $\sigma_{f_*|y}^2$ can be expressed in a similar way. The expressions for variance are omitted also in the following sections since their derivation is similar to the ones of the posterior mean [3, ch. 2] [4, sec. 2.3].

IV. GPR Under Uncertain Training Positions

In the previous section the training positions $\boldsymbol{x}^{(i)}$ were considered known. In practical applications this information may not be available. We shall therefore consider that the training positions are known only up to their probability density functions (pdfs) $p(\boldsymbol{x}^{(i)})$, which can be expressed for all the training positions jointly as

$$p(\boldsymbol{x}_\mathrm{t}) = p(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(I)}) \ . \tag{16}$$

Since now the training positions are random and the GP RVs $\boldsymbol{f}$ need to have their position explicitly specified by conditioning on the random training positions to attain the original meaning, i.e.,

$$\boldsymbol{f}|\boldsymbol{x}_\mathrm{t} = (f(\boldsymbol{x}^{(1)}) \quad f(\boldsymbol{x}^{(2)}) \quad \ldots \quad f(\boldsymbol{x}^{(I)}))^\mathrm{T} \in \mathbb{R}^I \ . \tag{17}$$

The posterior distribution as in (10) is then obtained by conditioning all the pdfs on $\boldsymbol{x}_\mathrm{t}$, which yields

$$p(f_*|\boldsymbol{y}, \boldsymbol{x}_\mathrm{t}) = \frac{p(f_*, \boldsymbol{y}|\boldsymbol{x}_\mathrm{t})}{p(\boldsymbol{y}|\boldsymbol{x}_\mathrm{t})} \ . \tag{18}$$

The term in (18) corresponds to the standard GPR setup with known training positions $\boldsymbol{x}_\mathrm{t}$ and would operate equally as (10) if we knew them. Since we do not know $\boldsymbol{x}_\mathrm{t}$ and only the pdf (16) of this vector is available, we need to express the *unconditional* pdf the GP RV at the test position f_* given the observations vector $\boldsymbol{y}$, i.e., $p(f_*|\boldsymbol{y})$, which can be obtained from $p(f_*, \boldsymbol{x}_\mathrm{t}|\boldsymbol{y})$ using the *sum rule* as

$$p(f_*|\boldsymbol{y}) = \int_{\mathbb{R}^{DI}} p(f_*, \boldsymbol{x}_\mathrm{t}|\boldsymbol{y}) \, d\boldsymbol{x}_\mathrm{t} \ , \tag{19}$$

which can be further developed using the *product rule* as

$$p(f_*|\boldsymbol{y}) = \int_{\mathbb{R}^{DI}} p(f_*|\boldsymbol{y}, \boldsymbol{x}_\mathrm{t}) p(\boldsymbol{x}_\mathrm{t}) \, d\boldsymbol{x}_\mathrm{t} \ . \tag{20}$$

The posterior pdf (20) is not a Gaussian pdf in general. Nevertheless, we will approximate it as a Gaussian pdf using its 1st two moments to achieve a similar term as in (14), i.e.,

$$p(f_*|\boldsymbol{y}) \sim \mathcal{N}(f_*;\ \mathrm{E}\{f_*|\boldsymbol{y}\},\ \mathrm{var}\{f_*|\boldsymbol{y}\}) \ . \tag{21}$$

A. Posterior Mean

The mean value of the posterior distribution in (21) can be expressed as

$$\mathrm{E}\{f_*|\boldsymbol{y}\} = \int_{\mathbb{R}} f_* p(f_*|\boldsymbol{y}) \, df_* \ , \tag{22}$$

which after plugging (20) becomes

$$\begin{aligned} \mathrm{E}\{f_*|\boldsymbol{y}\} &= \int_{\mathbb{R}} f_* \int_{\mathbb{R}^{DI}} p(f_*|\boldsymbol{y}, \boldsymbol{x}_\mathrm{t}) p(\boldsymbol{x}_\mathrm{t}) \, d\boldsymbol{x}_\mathrm{t} \, df_* && (23) \\ &= \int_{\mathbb{R}^{DI}} \int_{\mathbb{R}} f_* p(f_*|\boldsymbol{y}, \boldsymbol{x}_\mathrm{t}) \, df_* p(\boldsymbol{x}_\mathrm{t}) \, d\boldsymbol{x}_\mathrm{t} && (24) \\ &= \int_{\mathbb{R}^{DI}} \mu_*(\boldsymbol{x}_\mathrm{t}) p(\boldsymbol{x}_\mathrm{t}) \, d\boldsymbol{x}_\mathrm{t} \ , && (25) \end{aligned}$$

where

$$\mu_*(\boldsymbol{x}_\mathrm{t}) \triangleq \mathrm{E}\{f_*|\boldsymbol{y}, \boldsymbol{x}_\mathrm{t}\} = \int_{\mathbb{R}} f_* p(f_*|\boldsymbol{y}, \boldsymbol{x}_\mathrm{t}) \, df_* \tag{26}$$

is the standard GPR posterior mean value according to (15) considering known training positions $\boldsymbol{x}_\mathrm{t}$.

V. Monte Carlo Approximation

The integrals used for evaluating the posterior mean in (25) cannot be computed in closed form in general. As suggested in [1], this problem can be solved by using the *Monte Carlo* (MC) approximation, which provides a solution in a computationally feasible way.

To express the MC approximation of the posterior mean in (25) we firstly reformulate it as an expectation with respect to the pdf $p(\boldsymbol{x}_\mathrm{t})$, i.e.,

$$\mathrm{E}\{f_*|\boldsymbol{y}\} = \mathrm{E}^{p(\boldsymbol{x}_\mathrm{t})}\{\mu_*(\boldsymbol{x}_\mathrm{t})\} \ . \tag{27}$$

The pdf of training positions $p(\boldsymbol{x}_\mathrm{t})$ can now be approximated using a set of samples $\boldsymbol{x}_{\mathrm{t},i}, i = 1, \ldots, s$ drawn according to $p(\boldsymbol{x}_\mathrm{t})$ with a point-mass function

$$p_\mathrm{MC}(\boldsymbol{x}_\mathrm{t}) = \frac{1}{s} \sum_{i=1}^{s} \delta(\boldsymbol{x}_\mathrm{t} - \boldsymbol{x}_{\mathrm{t},i}) \ , \tag{28}$$

where $\delta(\cdot)$ is the Dirac delta function. The posterior mean in (27) can then be approximated as

$$\begin{aligned} \mathrm{E}^\mathrm{MC}\{f_*|\boldsymbol{y}\} &= \mathrm{E}^{p_\mathrm{MC}(\boldsymbol{x}_\mathrm{t})}\{\mu_*(\boldsymbol{x}_\mathrm{t})\} \\ &= \frac{1}{s} \sum_{i=1}^{s} \mu_*(\boldsymbol{x}_{\mathrm{t},i}) \ . && (29) \end{aligned}$$

It shall be noted that the computation complexity of (29) grows linearly with the number of MC samples s compared to the standard GPR as in (15).

VI. Simulation Setup

To verify the improved performance of the MC-GPR we developed a modular simulation setup in Python available at [5]. A block diagram representing simplified functionality of a single simulation run is depicted in Figure 2.

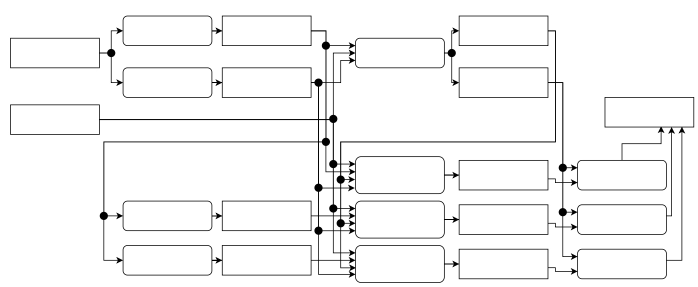


Fig. 2. Simulation setup scheme.

In order to be able to evaluate the performance of the investigated method in a fully described case, several simulation parameters need to be assumed. We shall consider a 2-dimensional space spanned by a rectangle of size 2×1 in arbitrary spatial units. Within this rectangle, 10 random, uniformly distributed training positions are assumed and sampled.

The unknown spatial function (GP) is sampled at the test positions jointly with training positions. The test positions are placed uniformly on a grid within the considered rectangle with density of 10 test positions per spatial unit, i.e., resulting in $21 \times 11 = 231$ test positions. The estimates of spatial function and root-mean-square error (RMSE) evaluation are further performed only on these test positions.

The mean function of the GP used for simulations was considered to be constant, zero. The covariance function was chosen to be in the squared exponential form according to

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma^2 \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|^2}{2\sigma_{\mathrm{x}}^2}\right) , \tag{30}$$

where the variance and spatial scale parameters were chosen as $\sigma^2 = 1$ and $\sigma_{\mathrm{x}}^2 = 1$ respectively. The result of sampling a GP is depicted in Figure 1, which represents the ground-truth values of a spatial function to be estimated later. The observations of the spatial function at the training positions were evaluated according to (6), where the measurement noise variance was chosen as $\sigma_\epsilon^2 = 10^{-4}$.

Evaluating the standard GPR with known training positions according to (15) results in Figure 3 for the posterior mean and in Figure 4 for the posterior variance respectively.

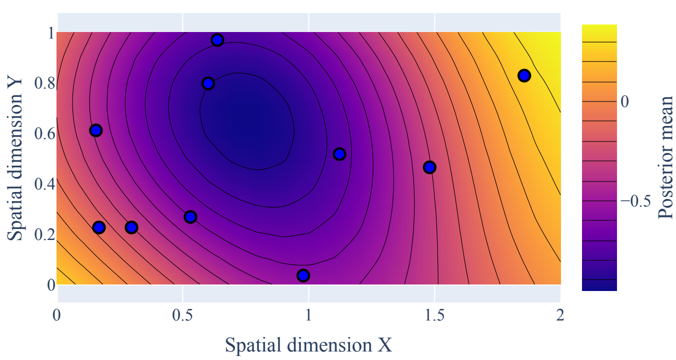


Fig. 3. GPR posterior mean values (represented by contour levels) using true training positions (indicated by blue dots).

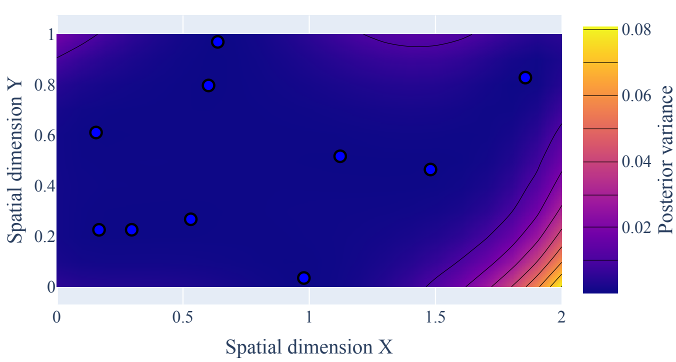


Fig. 4. GPR posterior variance (represented by contour levels) using true training positions (indicated by blue dots).

Since the training positions shall be further assumed uncertain, we will consider that only observations $\tilde{\boldsymbol{x}}^{(i)}$ containing additive Gaussian noise according to

$$p(\tilde{\boldsymbol{x}}^{(i)}) = \mathcal{N}(\tilde{\boldsymbol{x}}^{(i)};\, \boldsymbol{x}^{(i)},\, \sigma_{\mathrm{v}}^2 \mathbf{I}_I) , \qquad i = 1, \dots, I . \tag{31}$$

with $\sigma_{\mathrm{v}}^2 = 10^{-2}$ are available. Using these observed training positions instead of the original ones in the standard GPR method results in degraded estimation quality, which is apparent in Figure 5 showing the posterior mean and in Figure 6 showing the posterior variance.

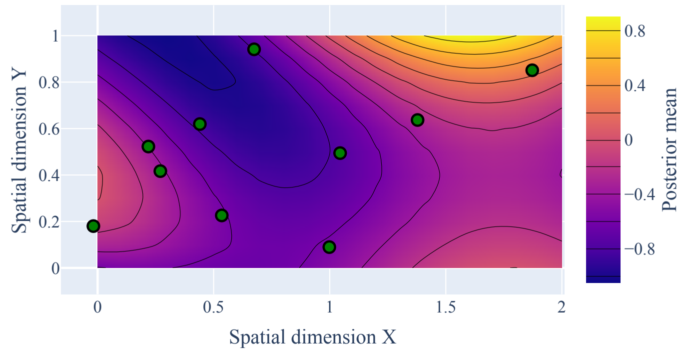


Fig. 5. GPR posterior mean values (represented by contour levels) using observed training positions (indicated by green dots).

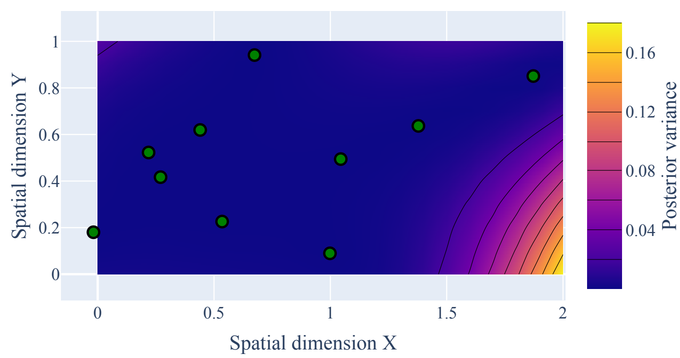


Fig. 6. GPR posterior variance (represented by contour levels) using observed training positions (indicated by green dots).

To account for the training positions uncertainty, we simulated the MC-GPR method according to (29) using $s = 100$ training positions samples generated according the posterior distribution of the training positions with uninformative prior. The resulting posterior mean is depicted in Figure 7 and posterior variance in Figure 8.

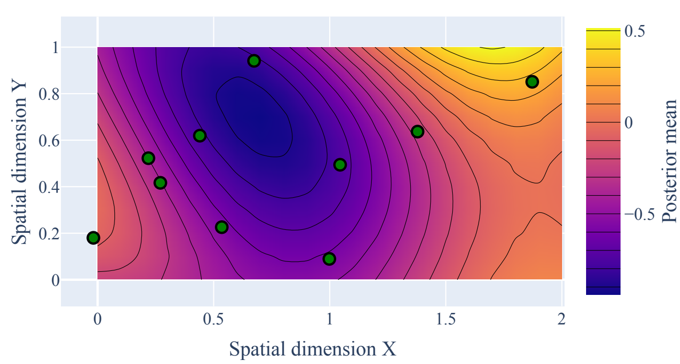


Fig. 7. MC-GPR posterior mean values (represented by contour levels) using observed training positions (indicated by green dots).

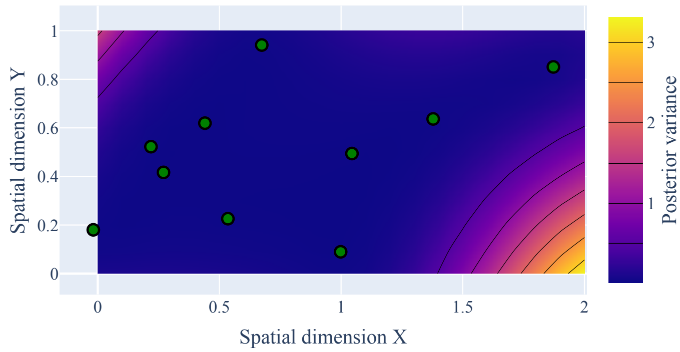


Fig. 8. MC-GPR posterior variance (represented by contour levels) using observed training positions (indicated by green dots).

VII. Results

As can be seen in the comparison of the true spatial function value in Figure 1 and its GPR estimate using the true training positions in Figure 3, this approach results in a high-quality estimation with low estimation uncertainty (variance), which can be seen in Figure 4. On the other hand, using the GPR approach in combination with uncertain (observed) training position without accounting for uncertainty results in a lower-quality estimate, as can be seen in Figure 5 combined with low estimation uncertainty (variance), which may be described as being unreasonably confident about the estimates. The MC-GPR approach accounting for uncertain training positions on the other hand resulted in better-quality estimate as can be seen in Figure 7 combined with higher estimation uncertainty (variance), which is a reasonable outcome of adding uncertainty to the inputs.

To quantitatively evaluate the estimation quality of the different approaches we performed 100 simulation runs, where each simulation run included sampling new GP realization, new training positions, new observation noises and new MC training positions samples. Evaluating the RMSE over all the simulation runs and all the test positions for different GPR methods resulted in Table I. Here we see that the best results were achieved by the standard GPR using known training positions. The worst results were attained by the GPR method using observed training positions naively. Compared to that, a significant performance improvement is achieved by the MC-GPR method, which accounts for training positions uncertainty via MC sampling.

TABLE I
RMSE of the considered GPR methods.

GPR method	RMSE
GPR, true training positions	0.1228
GPR, observed training positions	0.4027
MC-GPR	0.2810

VIII. Conclusion

We presented a derivation of the GPR operating on uncertain training positions. Having obtained an expression without closed-form solution, we presented its approximation using MC samples, which results in the MC-GPR method, which accounts for uncertain training positions. We further developed a simulation environment to compare the analyzed spatial function estimation approaches and to verify the performance improvements. In the individual simulation runs, it was demonstrated that MC-GPR method provides posterior distribution with higher variance than the usage of standard GPR with uncertain inputs, which corresponds to the increased uncertainty of the provided estimates and shows that the standard GPR is unreasonably confident about the provided estimates. Moreover, evaluating a higher number of simulation runs in the selected scenario setup and comparing the average RMSE of the estimates shows that the MC-GPR method achieves a significant improvement over the standard GPR operating on uncertain training positions. On the other hand, the performance of MC-GPR is still inferior to the standard GPR operating on known training positions and the performance gain comes at the cost of increased computation complexity linear with the number of MC samples. This makes the GPR with uncertain training positions an open problem with plausible improvements, which may be inspired by [6], [7], [2], [8].

References

[1] M. Jadaliha, Y. Xu, J. Choi, N. S. Johnson, and W. Li, "Gaussian Process Regression for Sensor Networks Under Localization Uncertainty," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 223–237, Jan. 2013.

[2] L. S. Muppirisetty, T. Svensson, and H. Wymeersch, "Spatial Wireless Channel Prediction under Location Uncertainty," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1031–1044, Feb. 2016.

[3] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning* (Adaptive computation and machine learning), en. Cambridge, Mass: MIT Press, 2006, OCLC: ocm61285753.

[4] C. M. Bishop, *Pattern recognition and machine learning* (Information science and statistics), en. New York: Springer, 2006.

[5] M. Ptáček, *Publications*, original-date: 2023-02-21T09:41:21Z, Mar. 2023. [Online]. Available: https://github.com/MarPta/Research (visited on 03/09/2023).

[6] A. Girard, "Approximate Methods for Propagation of Uncertainty with Gaussian Process Models," en, p. 167,

[7] D. Cervone and N. S. Pillai, "Gaussian Process Regression with Location Errors," *arXiv:1506.08256 [math, stat]*, Jun. 2015, arXiv: 1506.08256. [Online]. Available: http://arxiv.org/abs/1506.08256 (visited on 03/08/2021).

[8] M. Fröhle, L. S. Muppirisetty, and H. Wymeersch, "Channel gain prediction for multi-agent networks in the presence of location uncertainty," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 3911–3915.

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

Beat Tracking: Is 44.1 kHz Really Needed?

1st Matěj Ištvánek
Department of Telecommunications
Brno University of Technology, Technicka 12
616 00 Brno, Czech Republic
matej.istvanek@vut.cz

2nd Štěpán Miklánek
Department of Telecommunications
Brno University of Technology, Technicka 12
616 00 Brno, Czech Republic
stepan.miklanek@vut.cz

***Abstract*—Beat tracking is essential in music information retrieval, with applications ranging from music analysis and automatic playlist generation to beat-synchronized effects. In recent years, deep learning methods, usually inspired by well-known architectures, outperformed other beat tracking algorithms. The current state-of-the-art offline beat tracking systems utilize temporal convolutional and recurrent networks. Most systems use an input sampling rate of 44.1 kHz. In this paper, we retrain multiple versions of state-of-the-art temporal convolutional networks with different input sampling rates while keeping the time resolution by changing the frame size parameter. Furthermore, we evaluate all models using standard metrics. As the main contribution, we show that decreasing the input audio recording sampling frequency up to 5 kHz preserves most of the accuracy, and in some cases, even slightly outperforms the standard approach.**

***Index Terms*—Beat tracking, music information retrieval, temporal convolutional networks, machine learning.**

I. Introduction

In Music Information Retrieval (MIR), one of the core tasks is beat tracking or beat detection. It aims at detecting "tactus" positions in an audio signal — described as "the most comfortable foot-tapping rate when unconsciously tapping to a piece of music" [1]. Early conventional approaches to beat tracking usually utilized a two-stage strategy. First, an onset detection function was computed from time-frequency representations, such as spectrograms or mel-spectrograms. Then a post-processing phase with prior musical knowledge was implemented to determine which onsets might correspond to beats. Well-known examples of non-machine learning approaches are BeatRoot [2], dynamic programming-based method [3], and Predominant Local Pulse [4], [5].

Over the years, numerous beat tracking methods have been proposed, ranging from rule-based systems and probability methods to machine learning models. With the rise of deep learning techniques, beat tracking has significantly shifted toward data-driven deep neural networks. The increasing availability of data and their annotation[1] helped to outperform every conventional non-machine learning beat tracker (see MIREX results[2]). However, expressive music and genres with more complex metric and rhythmic structures are still considered highly challenging. The Recurrent Neural Networks (RNN) proved that data-driven approaches provide better results than conventional systems [6], [7]. Furthermore, the multi-model approach [8] was implemented to reflect different rhythms depending on the music style and genre. Selecting the best-performing model as the state-of-the-art is challenging, due to the absence of a beat tracking competition (the last MIREX beat tracking competition ended in 2019, and further evaluation is based on the beat tracking community). However, many studies use *n*-fold cross-validation and similar datasets, achieving more than 90 % F-score (explained in Section III-C) for non-classical and less expressive music.

The Temporal Convolutional Networks (TCNs), based on the original implementation of WaveNet [9], are one of the newest methods for beat tracking and usually achieve the highest F-score. The well-known examples are [10] and [1]. All mentioned models use time-frequency representation (modified spectrograms). Authors in [11] show an end-to-end approach using time-domain representation, feeding raw audio samples into the TCN, achieving similar results as state-of-the-art systems. It is also possible to train beat, downbeat, and tempo activation functions jointly [1], [12], solving more tasks with just one model. For more details about deep learning beat and downbeat tracking TCN models, we refer to [13].

In this paper, we implement multiple TCN beat tracking systems and evaluate their abilities to detect beats on standard datasets. We add skip connections to the networks as one of the possible network modifications and treat them as separate models. As the main contribution, we train five models on 44.1, 22.05, 11.025, and 5.5 kHz input sampling rates to demonstrate how much higher-frequency information is needed for the beat tracking task. We show that lower sampling rates slightly outperform the standard approach with 44.1 kHz input signal in most models. This may be useful for applying beat tracking systems in other MIR-related tasks, such as improving the synchronization accuracy when used jointly with Dynamic Time Warping methods [14] without the need for resampling. The results indicate that even a system trained on the 5.5 kHz input audio signal is comparable to the standard 44.1 kHz model. The higher frequency content seems redundant for the universal beat tracking task.

The rest of the paper is organized as follows. Section II describes the architecture, pre-processing, and models of TCN beat tracking systems. Section III explains the training and evaluation process. Finally, Section IV shows the results followed by discussion and conclusions in Section V.

[1]Beat annotation consists of discrete time points of beat occurrence.
[2]https://www.music-ir.org/mirex/wiki/2019:Audio_Beat_Tracking (accessed on 27 March 2023)

DOI 10.13164/eeict.2023.227

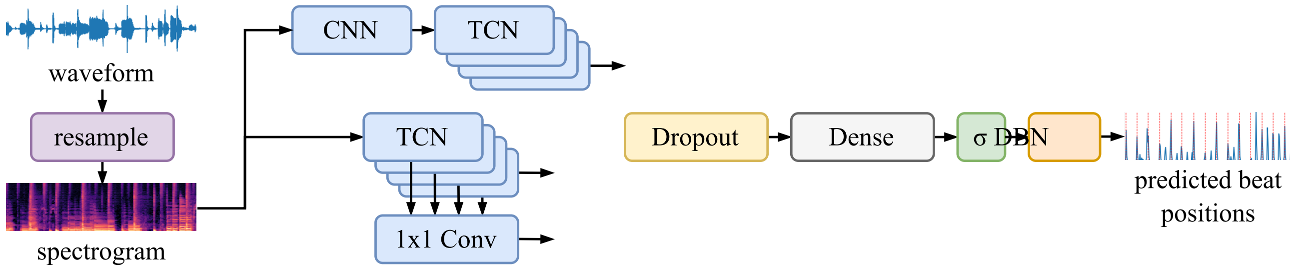


Fig. 1. High-level overview of different approaches to TCN beat tracking.

II. Methods

A. Architecture

Temporal convolutional networks are a class of deep neural networks that have gained popularity in various time-series applications. TCNs consist of multiple layers of temporal convolutions and non-linear activations, allowing them to capture long-term dependencies in sequential data. The ability to model long-term dependencies makes TCNs an attractive option for beat tracking, as the tempo of a musical piece is inherently a temporal pattern, with the exception of expressive performances. For beat tracking, the input to the TCN is a sequence of audio features, usually modified spectrograms. The output of the TCN is a sequence of beat activations, representing the likelihood of a beat occurring at each time step. The beat activations are then post-processed to estimate the exact time positions of beats. We apply a post-processing method called Dynamic Bayesian Network (DBN) [15] as a standard approach to obtain a sequence of beats.

To adapt TCNs to beat tracking, researchers have proposed various modifications to the standard TCN architecture. One of the modifications is to use skip connections, allowing the network to bypass certain layers and directly propagate information from earlier to later ones. Skip connections have been shown to improve the training stability and convergence of TCNs. In our paper, we experiment with three slightly different versions of the TCN beat tracker and modify two of them with additional skip connections.

Figure 1 shows different variants of beat tracking neural networks. The first approach is to use a two-dimensional Convolutional Neural Network (CNN) to extract musically motivated features from the spectrograms and then use a sequence of TCN blocks to capture temporal information. In this work, we experiment with discarding the CNN and using only the TCN blocks to reduce the number of trainable parameters. We also utilize skip connections and combine the intermediate outputs of the TCN blocks using a 1×1 convolutional layer instead of taking only the output from the last TCN block. We evaluate all models using 44.1, 22.05, 11.025, and 5.5 kHz as input sampling rates.

B. Pre-processing

We use frame sizes 2048, 1024, 512, and 256 samples to maintain the same temporal context for 44.1, 22.05, 11.025, and 5.5 kHz sampling rates, respectively. However, the 5.5 kHz sampling rate is a rounded number (the correct rate would be 5 512.5, which is impractical). Therefore, this model does not exactly follow the sampling rate/frame size compromise. We apply the Short-Time Fourier Transform to the audio frames, followed by a filter bank to obtain magnitude spectrograms with logarithmically spaced frequency bins. We refer to [13] for a detailed description of the pre-processing step. The number of frequency bins per octave was set to 12. We ensured that the fps = 100 stayed the same for each scenario. The time resolution also corresponds to the output beat activation function.

C. Models

We use the model `bock_2020` from [1] as a baseline for our experiments. This model includes a CNN to extract relevant spectral features from the spectrograms, which are then fed as input to the first TCN block. The inner structure of the TCN block is shown in Figure 2. The input gets first processed by two parallel dilated convolutional layers with different dilation rates. The output of the layers is then concatenated by the channel dimension, followed by an Exponential Linear Unit (ELU) activation function. The next block is a spatial dropout layer used only during training to prevent overfitting. We set the value of spatial dropout to 0.1 in all experiments. Then, a 1×1 convolutional layer is used to reduce the number of convolution channels in half. The TCN block also contains a residual connection with an additional 1×1 convolutional layer, which helps to retain information from previous TCN blocks.

We also use a simplified TCN block described in [16] and implemented in `simple_tcn` and `tcn_dp`. The structure is depicted in Figure 3. The models are listed below; each row represents different architecture:

- `bock_2020_`x
- `simple_tcn_`x
- `simple_tcn_skip_`x
- `tcn_dp_`x
- `tcn_dp_skip_`x

To differentiate between various inputs of each model, we added a postfix: x stands for 44, 22, 11, or 5, which is equal to 44.1, 22.05, 11.025, and 5.5 kHz sampling rates, respectively.

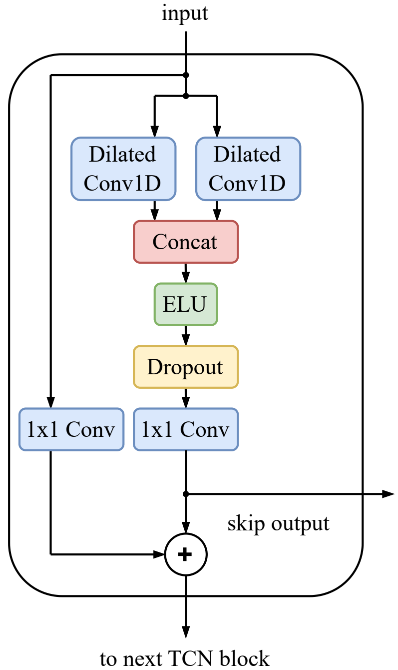


Fig. 2. Diagram of a TCN block used in [1].

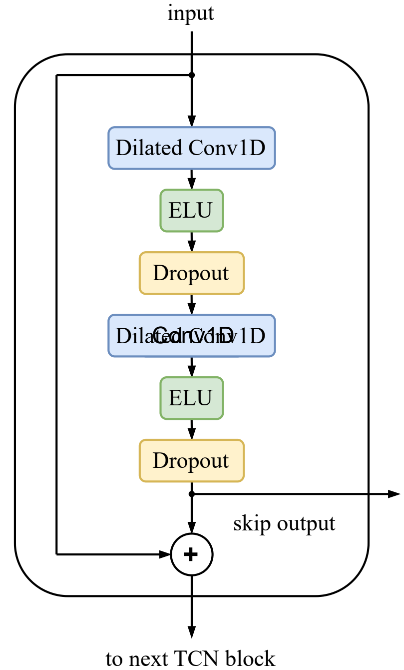


Fig. 3. Diagram of a TCN block presented in [16].

III. Experiments

A. Dataset

In our experiments, we use well-known datasets that have been used for beat tracking tasks for many years. We used the corrected annotations from S. Böck[3] with the exception of the Beatles dataset, in which all annotations were manually corrected to the corresponding ground-truth beat positions based on [16]. All datasets combined consist of 2 263 recordings with a total duration of around 26 hours and 175 127 ground-truth beat annotations. The list of datasets is described below:

- Ballroom [17] – excerpts around 30 s in length, dance music genres such as cha-cha, jive, quickstep, rumba, waltz, or tango,
- Hainsworth [18] – excerpts around 60 s in length, organized into six categories: rock/pop, dance, jazz, folk, classical, and choral music,
- GTZAN [19] – a large dataset containing 30 s excerpts and 10 different genres,
- SMC [20] – excerpts around 40 s in length, specifically selected to be challenging for the state-of-the-art beat tracking systems (for example, expressive performances, local tempo deviations, or complex music compositions),
- Beatles [21] – a collection of songs from Beatles with corrected annotations based on [1] and [16].

B. Training

First, we merge all datasets from Section III-A into one dataset and split it to train, test, and validation sets using the 80/10/10 strategy, respectively. Train and validation sets are shuffled for training, but the test set always contains the same recordings and annotation data.

We train each model on the training data while monitoring the performance on the validation set. We use the following settings: Adam optimizer, binary cross-entropy loss, and reduction of learning rate by a factor of 0.2 if the training does not improve for 10 epochs with the lower bound of learning rate set to 1×10^{-7}. Furthermore, early stopping is called if the change of validation loss is less than 1×10^{-4} for more than 20 epochs. The best checkpoint is then saved as the final model. Contrary to the original implementations, we use an augmentation inspired by [11]. During training, we shift the beat positions forward or back by a random amount between ± 70 ms. Table I shows the number of trainable parameters and training time for all models. The average training time of all proposed models combined was 50 minutes. The trainable parameters of the networks ranged from 48 481 to 71 521.

The model `bock_2020` derived from [1] was trained only on 44.1 and 22.05 kHz sampling rates, and without skipping modifications. Changing the network's input size was impossible without changing the inner structure — for example, convolution channels or the dilation factor. However, we decided to keep it in our experiments and show the difference between the original 44.1 and 22.05 kHz models.

The training and validation losses are shown in Figures 4 and Figure 5, respectively. We only display one of the models (`tcn_dp_skip`) with all sampling rates for brevity.

[3]https://github.com/superbock/ISMIR2020 (accessed on 27 March 2023)

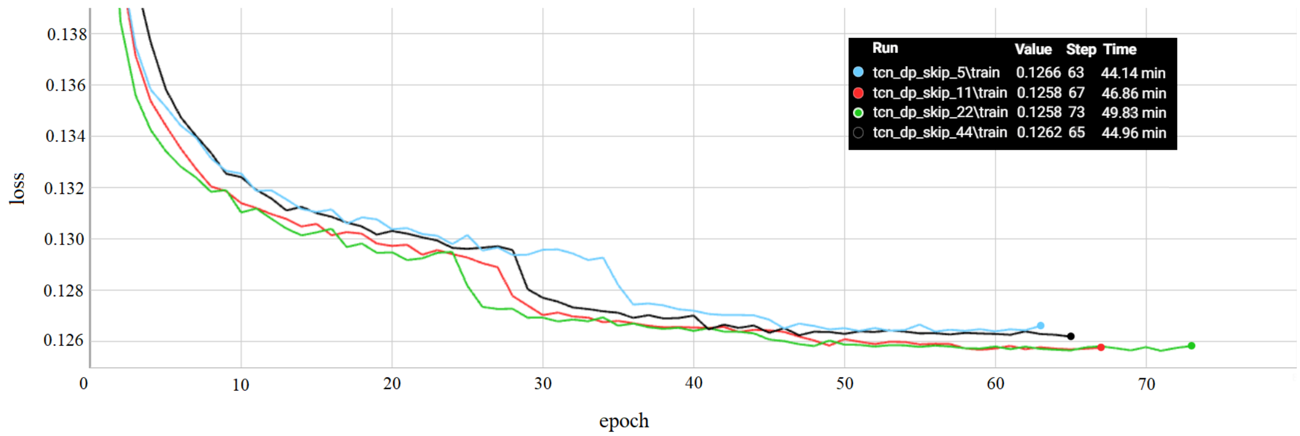


Fig. 4. Loss of the `tcn_dp_skip` model on the training data for each epoch.

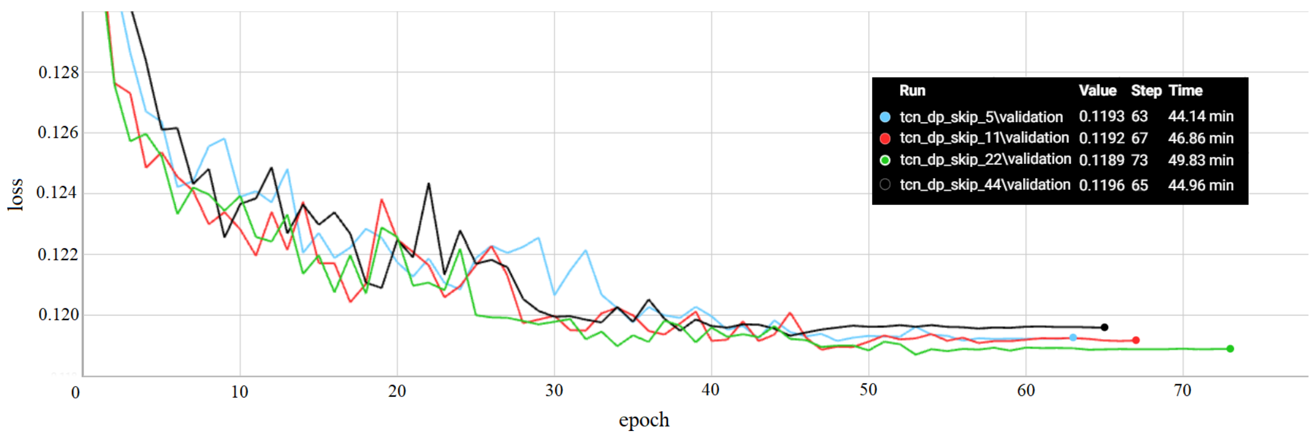


Fig. 5. Loss of the `tcn_dp_skip` model on the validation data for each epoch.

C. Evaluation

We use standard F-score metrics on the test set to evaluate proposed models in terms of prediction accuracy. The F-score is a harmonic mean of precision and recall based on true positives, false positives, and false negatives [22]. To decide if the target beat is within the range of ground-truth beat position, we use a window of length 70 ms, which is a standard value in the beat tracking community [23]. We compute additional metrics (Cemgil, P-score, Goto, and CMLc metrics) and refer to [21] for more details about their implementation. Contrary to other studies, we do not use *n*-fold cross-validation due to the nature of our experiments.

IV. Results

We evaluated all models on the test set described in Section III-A and III-B. Table II shows the F-score, Cemgil, P-score, Goto, and CMLc metrics for each architecture and sampling rate. Bold numbers indicate the best result for given metrics and architecture. The `bock_2020_22` model achieves the highest scores overall (F-score = 0.928, Cemgil = 0.829, P-score = 0.912, Goto = 0.828, and CMLc = 0.840), surpassing the 44.1 kHz version. Lower sampling rates slightly increase the models' accuracy. An exception is the `tcn_dp_44` model with F-score = 0.912 compared to the `tcn_dp_22` model with F-score = 0.900. The differences, however, are not significant. Furthermore, `tcn_dp_skip_22` is comparable to the state-of-the-art beat tracking model `bock_2020_44` with worse Cemgil and slightly better P-score and CMLc metrics. The difference between the training and validation process of the same architecture but varied input sampling rates is shown in Figures 4 and 5. There is no connection between training time and the input sampling frequency due to the early stopping mechanism.

V. Discussion and Conclusions

In this paper, we trained multiple beat tracking systems with slightly modified architectures on standard datasets and evaluated their performance. Using additional skip connections increased the metrics in most cases, except for `tcn_dp_44` and `simple_tcn_22` models. The well-known `bock_2020` system achieved the highest detection accuracy when trained on a 22.05 kHz audio input sampling rate, although its authors used 44.1 kHz. All networks except `tcn_dp` provided better results when trained on lower sampling rates. This may be thanks

TABLE I

THE NUMBER OF PARAMETERS, TRAIN TIME IN SECONDS FOR EACH MODEL, AND THE MEAN TRAIN TIME FOR EACH ARCHITECTURE.

model	params	train time [s]	mean [s]
bock_2020_22	65 941	3 286	3 099
bock_2020_44	65 941	2 911	
simple_tcn_5	64 701	3 308	3 249
simple_tcn_11	67 341	2 782	
simple_tcn_22	69 981	3 635	
simple_tcn_44	71 521	3 272	
simple_tcn_skip_5	64 483	3 355	3 340
simple_tcn_skip_11	67 123	3 422	
simple_tcn_skip_22	69 763	3 474	
simple_tcn_skip_44	71 303	3 110	
tcn_dp_5	48 481	2 374	2443
tcn_dp_11	49 921	2 180	
tcn_dp_22	51 361	2 570	
tcn_dp_44	52 201	2 648	
tcn_dp_skip_5	48 683	2 777	2 918
tcn_dp_skip_11	50 123	2 951	
tcn_dp_skip_22	51 563	3 117	
tcn_dp_skip_44	52 403	2 826	

TABLE II

BEAT TRACKING EVALUATION OF ALL MODELS ON THE TEST SET USING STANDARD METRICS. BOLD NUMBERS INDICATE THE BEST RESULT FOR GIVEN METRICS AND ARCHITECTURE.

model	F-score	Cemgil	P-score	Goto	CMLc
bock_2020_22	**0.928**	**0.829**	**0.912**	**0.828**	**0.840**
bock_2020_44	0.925	0.815	0.904	0.806	0.821
simple_tcn_5	0.907	**0.799**	0.889	0.771	0.799
simple_tcn_11	0.900	0.773	0.878	0.744	0.778
simple_tcn_22	**0.917**	**0.799**	**0.903**	**0.789**	**0.823**
simple_tcn_44	0.907	0.765	0.887	0.767	0.798
simple_tcn_skip_5	0.907	0.772	0.890	**0.784**	0.810
simple_tcn_skip_11	**0.915**	**0.778**	**0.894**	0.771	0.801
simple_tcn_skip_22	0.907	0.767	0.890	0.775	0.807
simple_tcn_skip_44	0.909	0.773	0.893	**0.784**	**0.811**
tcn_dp_5	0.899	0.790	0.879	0.767	0.783
tcn_dp_11	0.885	0.804	0.863	0.740	0.763
tcn_dp_22	0.900	**0.805**	0.884	0.758	0.796
tcn_dp_44	**0.912**	**0.805**	**0.896**	**0.806**	**0.813**
tcn_dp_skip_5	0.905	0.751	0.890	0.771	0.799
tcn_dp_skip_11	0.918	0.770	0.900	0.784	0.809
tcn_dp_skip_22	**0.925**	**0.776**	**0.912**	**0.806**	**0.838**
tcn_dp_skip_44	0.909	0.764	0.895	0.775	0.806

to, for example, redundant information in higher frequencies. In most music genres, the beat structure is defined by lower frequencies and specific pulsations. Even 5.5 kHz models show comparable performance, considering many instruments contain overtones and timbre above 2.5 kHz. It seems that 44.1 kHz might not be needed for the beat tracking task. For some applications, the lower input sampling rates may be beneficial, as most of the common music processing pipelines and extraction tools work with 22.05 kHz audio signals. In the future, we want to build on these experiments and release open-source models with different input sampling rates to provide more options for subsequent applications.

REFERENCES

[1] S. Böck, J. S. Cardoso, and M. E. P. Davies, "Deconstruct, analyse, reconstruct: How to improve tempo, beat, and downbeat estimation," in *21st International Society for Music Information Retrieval Conference (ISMIR)*, 2020.

[2] S. Dixon, "An interactive beat tracking and visualisation system," in *Proceedings of the 2001 International Computer Music Conference (ICMC'2001)*, 2001.

[3] D. P. Ellis, "Beat tracking by dynamic programming," *Journal of New Music Research*, vol. 36, no. 1, pp. 51–60, 2007.

[4] P. Grosche and M. Müller, "Extracting predominant local pulse information from music recordings," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 6, pp. 1688–1701, 2011.

[5] ——, "Tempogram toolbox: Matlab implementations for tempo and pulse analysis of music recordings," in *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR)*, 2011.

[6] S. Böck and M. Schedl, "Enhanced beat tracking with context-aware neural networks," in *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx)*, 2011.

[7] S. Böck, F. Krebs, and G. Widmer, "Accurate tempo estimation based on recurrent neural networks and resonating comb filters," in *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, 2015.

[8] ——, "A multi-model approach to beat tracking considering heterogeneous music styles," in *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR 2014)*, 2014.

[9] A. van den Oord *et al.*, "WaveNet: A generative model for raw audio," *arXiv preprint arXiv:1609.03499v2*, 2016.

[10] M. E. P. Davies and S. Böck, "Temporal convolutional networks for musical audio beat tracking," in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019.

[11] C. J. Steinmetz and J. D. Reiss, "Wavebeat: End-to-end beat and downbeat tracking in the time domain," in *151st Audio Engineering Society Convention*, 2021.

[12] S. Böck, F. Krebs, and G. Widmer, "Joint beat and downbeat tracking with recurrent neural networks," in *17th International Society for Music Information Retrieval Conference (ISMIR)*, 2016.

[13] M. E. P. Davies, S. Böck, and M. Fuentes, *Tempo, Beat and Downbeat Estimation*. https://tempobeatdownbeat.github.io/tutorial/intro.html, November 2021, (accessed on 27 March 2023).

[14] Y. Özer, M. Ištvánek, V. Arifi-Müller, and M. Müller, "Using activation functions for improving measure-level audio synchronization," in *International Society for Music Information Retrieval Conference (ISMIR)*, 2022.

[15] F. Krebs, S. Böck, and G. Widmer, "An efficient state-space model for joint tempo and meter tracking," in *International Society for Music Information Retrieval Conference (ISMIR)*, 2015.

[16] T. Suchánek, "Detektor tempa hudebních nahrávek na bázi neuronové sítě," Master Thesis, Brno University of Technology, Department of Telecommunications, 2021.

[17] F. Krebs, S. Böck, and G. Widmer, "Rhythmic pattern modeling for beat and downbeat tracking in musical audio," in *International Society for Music Information Retrieval Conference (ISMIR)*, 2013.

[18] S. W. Hainsworth and M. D. Macleod, "Particle filtering applied to musical tempo tracking," *EURASIP Journal on Advances in Signal Processing*, vol. 15, 2004.

[19] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.

[20] A. Holzapfel, M. E. P. Davies, J. R. Zapata, J. L. Oliveira, and F. Gouyon, "Selective sampling for beat tracking evaluation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 9, pp. 2539–2548, 2012.

[21] M. E. Davies, N. Degara, and M. D. Plumbley, "Evaluation methods for musical audio beat tracking algorithms," Queen Mary University of London, Centre for Digital Music, Tech. Rep., 2009.

[22] M. F. McKinney, D. Moelants, M. E. P. Davies, and A. Klapuri, "Evaluation of audio beat tracking and music tempo extraction algorithms," *Journal of New Music Research*, vol. 36, no. 1, pp. 1–16, 2007.

[23] C. Raffel, B. McFee, E. J. Humphrey, J. Salamon, O. Nieto, D. Liang, D. P. Ellis, and C. C. Raffel, "Mir_eval: A transparent implementation of common mir metrics." in *15th International Society for Music Information Retrieval Conference (ISMIR)*, 2014.

Digital Biomarkers for Assessing Respiratory Disorders in Parkinson's Disease

1st Daniel Kováč
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xkovac41@vut.cz

2nd Dominik Cvetler
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
195798@vut.cz

***Abstract*—Respiratory disorders are a significant part of hypokinetic dysarthria (HD) that affects patients with Parkinson's disease (PD). Still, their potential role in the objective assessment of HD has not yet been fully explored, which is the primary goal of this study. Several respiratory features were designed and extracted from acoustic signals recorded during text reading. Based on these features, the XGBoost model was able to predict clinical test scores of phonorespiration with an estimated error rate of 12.54%. Statistical analysis revealed that measuring respiration rate and quantifying signal fluctuations during inspiration have great potential in the objective assessment of respiratory disorders in patients with PD.**

***Index Terms*—respiration, digital biomarkers, hypokinetic dysarthria, Parkinson's disease, statistics, machine learning**

I. Introduction

Despite being described over two centuries ago [1], the exact causes of Parkinson's disease (PD), a chronic neurodegenerative disorder [2]), remain unknown. While genetic predisposition and age are risk factors, exposure to pesticides, high caloric intake, or head injuries may also increase the likelihood of developing PD [3]. Unfortunately, a cure for PD is not yet available, but medication, such as levodopa, or more invasive treatments, such as deep brain stimulation, can help alleviate its symptoms [4] [5]. Additionally, the incidence of PD appears to be on the rise [6]. Therefore, early diagnosis plays a crucial role in managing the disease. Having tools to identify and monitor PD in its early stages is vital, as this can improve a patient's quality of life.

The initial signs of Parkinson's disease (PD) often involve speech and voice problems [7], which are collectively known as hypokinetic dysarthria (HD) [7]. People with HD may suffer from deteriorated respiration, phonation, articulation, resonances, timing and prosody [8]. While some PD patients may only experience speech difficulties in certain areas [9], up to 90% of cases involve impairment in at least one domain [10]. Common symptoms of HD include hoarseness [11], tremor [12], audible breath [13], hypernasality [14], inappropriate silences [15] and speech disfluency [16]. Additionally, speech may be quiet [17] and difficult to understand [18], often characterized by monopitch and monoloudness [19].

Supported by EU (project no. LX22NPO5107) and by the Czech Ministry of Health (grant no. NU22J-04-00074). Many thanks to Dr. Mekyska, head of the Brain Diseases Analysis Laboratory (BDALab).

Ergo, acoustic analysis of speech and voice is considered a valuable and practical tool for detecting and monitoring Parkinson's disease (PD) objectively. This analysis involves recording the patient's speech signal and processing it digitally to obtain features that quantify specific disorders. Comparing these features with those of healthy controls (HC) using statistical methods makes it possible to assess the severity of hypokinetic dysarthria (HD) in PD patients. Furthermore, by utilizing machine learning techniques, it becomes possible to predict the scores of tests administered by neurologists or clinical speech therapists. This can lead to a faster and more objective assessment of the patient's condition.

Although many studies have been on quantifying disorders linked with HD using recorded acoustic speech/voice signals, there is only one known aiming for the domain of respiration. Hlavnička et al. [20] analysed the connected speech and compared automatically extracted features of healthy controls, patients with rapid eye movement sleep behaviour disorder (RBD), and newly diagnosed PD patients. For measuring the respiratory deficits, they used relative loudness of respiration, the latency of respiratory exchange, pause intervals per respiration and rate of speech respiration. Repeated measures analysis of variance (RM-ANOVA) revealed no differences in HC and PD groups regarding these features.

Yet respiratory disorders are found a significant part of HD [21], and we should pay attention to it if we want to investigate all its dimensions [22]; moreover, they are the primary cause of mortality in PD [23]. Impaired respiratory muscle activity affects the ability to generate and control airflow during a speech, resulting in shallow and insufficient breaths. This can lead to a weak, breathy voice and reduced loudness. The symptoms may follow obstructive and restrictive ventilatory defect patterns [24]. Patients with PD may suffer from dyspnea, tachypnea, hyperventilation or apnea [25]. Other symptoms are hypophonia, shaky voice, stridor or wheeze [26].

The main objective of this work is to propose digital biomarkers quantifying respiratory disorders and fill the gap in the objective assessment of HD in this domain. New features will be designed, analysed and statistically tested. Clinical scores will be predicted based on these features using machine learning techniques, and finally, we will discuss the features' potential.

II. Method

A. Dataset

The recordings used in this study come from the PARCZ database [27]. To observe respiratory difficulties, we analyzed the signal recorded during text reading. The average duration of the reading task is 61.7 seconds. Table I contains the number of PD patients and HC, while their clinical data is presented in Table II. The mean age of HC and PD patients is 63.1 and 67.4 years, respectively.

TABLE I
DEMOGRAPHIC DATA

	female	male	total
HC	24	22	46
PD	36	57	93
total	60	79	139

TABLE II
CLINICAL DATA

	HC	PD
PD duration [year]	-	14.1 ± 2.8
LED [mg]	-	995.3 ± 551.1
faciokinesis	27.9 ± 1.8	24.6 ± 3.4
phonorespiration	28.5 ± 1.5	24.1 ± 3.6
phonetics	29.5 ± 1.0	25.7 ± 3.7
DX index	85.8 ± 3.2	74.4 ± 9.0
MMSE	28.3 ± 1.5	28.1 ± 2.3

[1] LED – Levodopa Equivalent Dose, DX index – Test 3F Dysarthric Profile (dysarthric index composed of faciokinesis, phonorespiration and phonetics) [28], MMSE – Mini Mental State Exam

B. Parametrization

We listened to the recordings and hand-marked each beginning and ending inspiration. We then used the PRAAT voice activity detector [29] to identify the speech parts of the signal. After that, we extracted the following features:

- **n_breath**: *The number of inspirations during the reading task.*
- **breath_rate**: *The number of inspirations relative to the total duration of the reading task.*
- **speech_length_mean**: *The mean duration of speech after inspiration.*
- **speech_length_std**: *The standard deviation of speech duration after inspiration relative to its mean.*
- **breath_length_mean**: *The mean duration of inspiration.*
- **breath_length_std**: *The standard deviation of inspiration duration relative to its mean.*
- **pause_before_mean**: *The mean pause between the end of the speech and the beginning of the inspiration.*
- **pause_before_std**: *The standard deviation of pause between the end of the speech and the beginning of the inspiration relative to its mean.*
- **pause_after_mean**: *The mean pause between the end of the inspiration and the beginning of the speech.*
- **pause_after_std**: *The standard deviation of pause between the end of the inspiration and the beginning of the speech relative to its mean.*
- **zcr_mean**: *The mean zero crossing rate of the inspiration.*
- **E_speech_std**: *The standard deviation of speech energy after inspiration relative to its mean.*
- **E_breath_std**: *The standard deviation of inspiration energy relative to its mean.*
- **E_breath_norm_mean**: *The mean inspiration energy relative to the mean speech energy.*

C. Statistical analysis

Initially, we used linear regression to remove the effect of age and gender. We then used the Shapiro-Wilk test to test the normal distribution of HC features. Next, we tried to determine whether the features of HC and PD patients come from the same probability distribution using the Mann-Whitney U test and calculated their mean, median and standard deviation. We created norms based on HC features and observed how many PD patients exceeded these norms. We set the lower T_{low} and upper T_{up} thresholds to:

$$T_{\text{low}} = Q_1 - 1.5 \cdot IQR, \tag{1}$$

$$T_{\text{up}} = Q_3 + 1.5 \cdot IQR, \tag{2}$$

where Q_1 and Q_3 are lower and upper quartiles and IQR is the interquartile range. Finally, we used Pearson's correlation coefficient to determine the relationship between features and clinical test scores.

D. Machine learning

We trained machine learning models using the XGBoost algorithm to predict clinical scores. Hyperparameters were tuned using random search, and the model was validated using 10-fold cross-validation with 20 repetitions. We observed the models' mean absolute error (MAE) and estimated error rate (EER) to measure its performance:

$$\text{EER} = \frac{\text{MAE}}{R}, \tag{3}$$

where R represents the range of values (of a clinical scale) in the training set.

III. Results

Table III shows the results of statistical tests, descriptive statistics and percentage ratio of PD patients outside the norms. We plotted the distribution of two features with the highest number of outliers in Fig. 1. Table IV then shows the results of the correlation tests. Table V presents the performance of the model's prediction, while Fig. 2 illustrates the graphical interpretation of the three most significant features during the prediction.

TABLE III
RESULTS OF STATISTICAL ANALYSIS

	Shapiro-Wilk p-value	Mann-Whitney p-value	$\bar{x}$ HC	$\bar{x}$ PD	$\tilde{x}$ HC	$\tilde{x}$ PD	σ HC	σ PD	Out of norm [%] under	Out of norm [%] above
n_breath	**0.053**	0.839	15.739	15.883	15.082	15.111	3.718	4.605	2.2	8.6
breath_rate	0.175	0.231	0.251	0.267	0.243	0.260	0.051	0.071	0	8.6
speech_length_mean	0.662	0.382	3.469	3.370	3.452	3.252	0.774	0.939	1.1	3.2
speech_length_std	**0.001**	0.891	0.471	0.477	0.447	0.432	0.126	0.184	0	4.3
breath_length_mean	0.160	0.481	0.389	0.383	0.379	0.372	0.061	0.080	1.1	2.2
breath_length_std	**0.009**	0.187	0.289	0.268	0.278	0.263	0.087	0.081	0	4.3
pause_before_mean	0.622	0.599	0.129	0.132	0.133	0.110	0.066	0.090	0	6.5
pause_before_std	**<0.001**	0.419	0.948	0.961	0.884	0.937	0.369	0.300	1.1	5.4
pause_after_mean	**0.001**	0.602	0.055	0.054	0.051	0.053	0.029	0.035	0	4.3
pause_after_std	**0.021**	0.659	1.172	1.143	1.033	1.086	0.456	0.52	0	1.1
zcr_mean	**<0.001**	0.577	0.022	0.026	0.018	0.020	0.013	0.022	0	10.8
E_speech_std	**<0.001**	0.333	0.928	0.845	0.772	0.727	0.551	0.496	0	4.3
E_breath_std	**<0.001**	0.260	1.395	1.291	1.211	1.129	0.766	0.916	0	6.5
E_breath_norm_mean	**<0.001**	0.195	0.050	0.043	0.038	0.032	0.041	0.042	0	6.5

[1] mean $\bar{x}$, median $\tilde{x}$ standard deviation σ
[2] highlighted p-values lower than 0.05

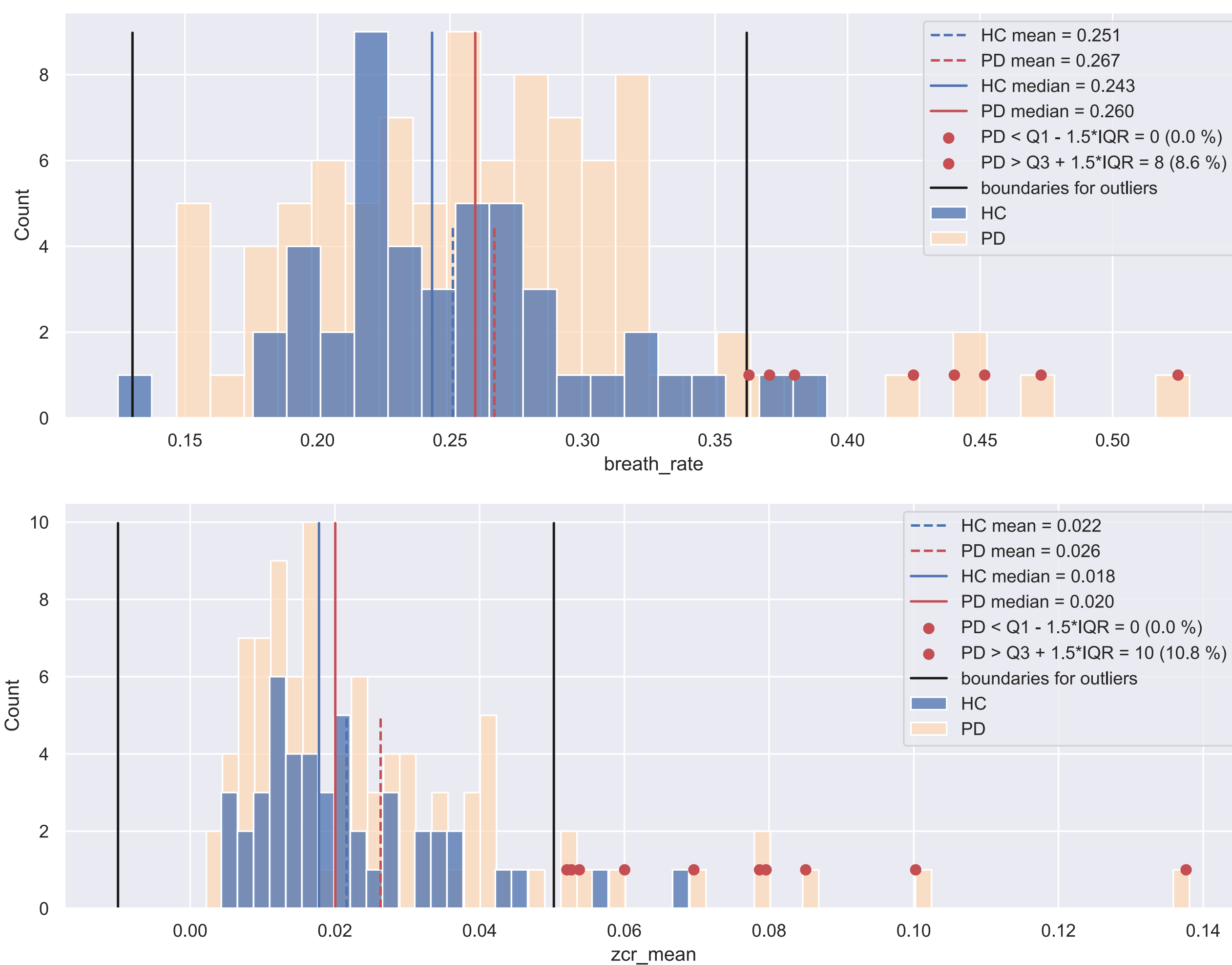


Fig. 1. Probability distributions of features with the highest number of PD patients outside the norm.

TABLE IV
PEARSON'S CORRELATION COEFFICIENTS

	faciokinesis		phonorespiration		phonetics		DX index		MMSE	
	coeff	p-value	coeff	p-value	coeff	p-value	coeff	p-value	coeff	p-value
n_breath	-0.027	0.750	-0.199	**0.019**	-0.066	0.437	-0.114	0.182	0.053	0.603
breath_rate	-0.026	0.766	-0.212	**0.012**	-0.060	0.483	-0.116	0.174	0.129	0.206
speech_length_mean	-0.078	0.360	0.091	0.287	-0.018	0.833	0.001	0.989	-0.201	**0.047**
speech_length_std	-0.026	0.760	-0.005	0.956	-0.059	0.491	-0.034	0.694	-0.106	0.301
breath_length_mean	-0.092	0.283	-0.103	0.226	-0.209	**0.013**	-0.154	0.071	-0.060	0.555
breath_length_std	0.005	0.950	-0.059	0.488	0.035	0.681	-0.008	0.924	-0.035	0.735
pause_before_mean	-0.072	0.396	-0.019	0.824	-0.112	0.191	-0.076	0.373	-0.011	0.913
pause_before_std	0.001	0.992	-0.016	0.849	0.078	0.361	0.024	0.782	0.093	0.364
pause_after_mean	-0.182	**0.032**	-0.151	0.076	-0.224	**0.008**	-0.210	**0.013**	-0.227	**0.025**
pause_after_std	-0.109	0.203	-0.072	0.397	0.081	0.345	-0.037	0.668	-0.005	0.964
zcr_mean	-0.195	**0.021**	-0.20	**0.018**	-0.126	0.140	-0.197	**0.020**	0.118	0.248
E_speech_std	-0.026	0.764	0.037	0.663	0.061	0.475	0.029	0.736	0.101	0.322
E_breath_std	-0.091	0.287	-0.054	0.529	-0.064	0.453	-0.078	0.359	0.139	0.173
E_breath_norm_mean	0.038	0.656	0.053	0.537	0.051	0.554	0.054	0.528	0.027	0.794

[1] highlighted p-values lower than 0.05

TABLE V
PERFORMANCE OF MACHINE LEARNING MODELS (MEAN AND STANDARD DEVIATION OF CROSS-VALIDATION)

	MAE	EER [%]
faciokinesis	2.72 ± 0.63	12.97 ± 3.00
phonorespiration	2.63 ± 0.70	12.54 ± 3.33
phonetics	2.98 ± 0.67	16.53 ± 3.69
DX index	6.53 ± 1.61	11.25 ± 2.77
MMSE	1.67 ± 0.60	11.91 ± 4.29

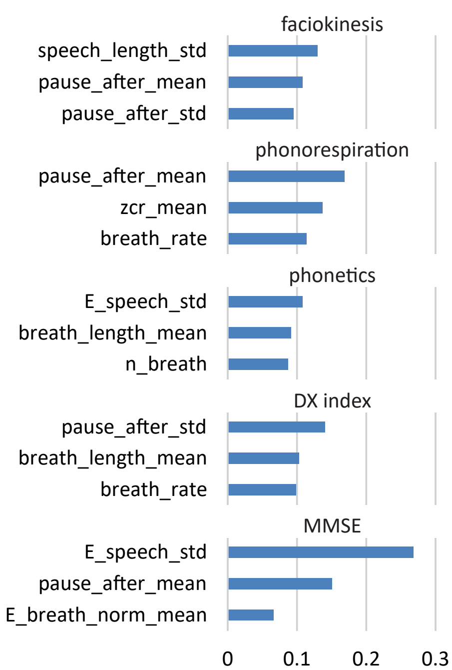


Fig. 2. Feature importances.

IV. DISCUSSION

We extracted 14 features based on knowledge of respiratory disorders in PD patients.

Statistical tests revealed that only 4 follow a normal distribution in the HC group (**breath_rate**, **speech_length_mean**, **breath_length_mean**, **pause_before_mean**), and thus PD patients are less likely to be out of the norm here for reasons other than HD.

No significant differences were detected between the features of HC and PD groups upon comparison, consistent with the study's results by Hlavnicka et al. [20]. However, we need to consider that the DX index of PD patients in our dataset is 74.4 on average, which is categorized as very mild dysarthria [28].

Nevertheless, there are several features with which many PD patients deviate from the HC norms. The best result in this respect shows the feature **zcr_mean**, which indicates that 10% of PD patients have higher signal fluctuation during inspiration, which may be related to stridor as it is defined as a high-pitched, harsh sound. Furthermore, 8% of patients had a higher need for breaths during reading (**breath_rate**), which may be a sign of tachypnea or hyperventilation (depending on shallow or deep breathing). Moreover, both features' mean, median and standard deviation are higher for people in the PD group.

Correlation tests show that there is a significant linear relationship between the features quantifying the number of breaths during reading (**n_breath**), the respiration rate (**breath_rate**) and the signal fluctuation during breathing (**zcr_mean**) and scores of a clinical test of phonorespiration. The features quantifying the average duration of the inspiration (**breath_length_mean**) and average pause between the inspiration and the beginning of the speech (**pause_after_mean**) correlate with the score of the phonetics test. It is apparent that respiration is closely linked to prosody, which is a large part of this test. An interesting finding is a strong negative correlation between features **pause_after_mean**, **speech_length_mean** and the Mini-Mental State exam results, meaning that people with cognitive deficits take longer to start speaking after inspiration and intervals between their breaths are longer.

Using machine learning models, we were able to predict scores on clinical scales to some extent. Specifically, based on the extracted features, we achieved an EER of 12.54% in predicting phonorespiration scores. Features **pause_after_mean**, **zcr_mean** and **breath_rate** were the most important during the prediction, underscoring their previously discussed significance.

A. Limitations

This work's limitation is a relatively small and unbalanced dataset. Furthermore, patients' speech was recorded when they were under the medication (ON state). As a result, the features extracted from the data may have been directly affected by the levodopa dose. Moreover, subjects were clinically tested just by one clinical speech therapist, and thus scores of the 3F test are given subjectively. The primary limitation of this study remains the manual annotation of the recordings, which is also subjective, time-consuming, and impractical for real-world applications. Therefore, as a follow-up to this research, implementing an automated system for detecting breaths in speech, re-extracting the features, and comparing the results with those obtained through manual annotation is proposed.

V. Conclusion

In this study, we analyzed several respiratory features extracted from the acoustic signal recorded during text reading. Based on statistical analysis and machine learning approaches, it was found that, in particular, measuring respiration rate and quantifying signal fluctuations during inspiration can be used to objectively assess respiratory disturbances associated with hypokinetic dysarthria in patients with Parkinson's disease.

References

[1] J. Parkinson, "An essay on the shaky palsy," *London: Sherwood, Neely and Jones*, pp. 1–6, 1817.

[2] O. Hornykiewicz, "Biochemical aspects of Parkinson's disease," *Neurology*, vol. 51, no. 2 Suppl 2, pp. S2–S9, 1998.

[3] A. Chade, M. Kasten, and C. Tanner, "Nongenetic causes of Parkinson's disease," *Parkinson's Disease and Related Disorders*, pp. 147–151, 2006.

[4] B. S. Connolly and A. E. Lang, "Pharmacological treatment of Parkinson disease: a review," *Jama*, vol. 311, no. 16, pp. 1670–1683, 2014.

[5] R. Baudouin, J. R. Lechien, L. Carpentier, J.-M. Gurruchaga, Q. Lisan, and S. Hans, "Deep brain stimulation impact on voice and speech quality in Parkinson's disease: A systematic review," *Otolaryngology–Head and Neck Surgery*, p. 01945998221120189, 2022.

[6] R. Savica, B. R. Grossardt, J. H. Bower, J. E. Ahlskog, and W. A. Rocca, "Time trends in the incidence of Parkinson disease," *JAMA neurology*, vol. 73, no. 8, pp. 981–989, 2016.

[7] D. B. Freed, *Motor speech disorders: diagnosis and treatment*. Plural Publishing, 2018.

[8] A. Rohl, S. Gutierrez, K. Johari, J. Greenlee, K. Tjaden, and A. Roberts, "Chapter 7 - speech dysfunction, cognition, and Parkinson's disease," in *Cognition in Parkinson's Disease*, ser. Progress in Brain Research, N. S. Narayanan and R. L. Albin, Eds. Elsevier, 2022, vol. 269, no. 1, pp. 153–173. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0079612322000176

[9] J. Rusz, T. Tykalova, M. Novotny, E. Ruzicka, and P. Dusek, "Distinct patterns of speech disorder in early-onset and late-onset de-novo Parkinson's disease," *npj Parkinson's Disease*, vol. 7, no. 1, pp. 1–8, 2021.

[10] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," *Behavioural neurology*, vol. 11, no. 3, pp. 131–137, 1998.

[11] M. Cernak, J. R. Orozco-Arroyave, F. Rudzicz, H. Christensen, J. C. Vásquez-Correa, and E. Nöth, "Characterisation of voice quality of Parkinson's disease using differential phonological posterior features," *Computer Speech & Language*, vol. 46, pp. 196–208, 2017.

[12] T. Tykalová, J. Rusz, J. Švihlík, S. Bancone, A. Spezia, and M. T. Pellecchia, "Speech disorder and vocal tremor in postural instability/gait difficulty and tremor dominant subtypes of Parkinson's disease," *Journal of Neural Transmission*, vol. 127, no. 9, pp. 1295–1304, 2020.

[13] Z. Thijs and C. R. Watts, "Perceptual characterization of voice quality in nonadvanced stages of Parkinson's disease," *Journal of Voice*, 2020.

[14] R. B. Hoodin and H. R. Gilbert, "Nasal airflows in parkinsonian speakers," *Journal of Communication Disorders*, vol. 22, no. 3, pp. 169–180, 1989.

[15] V. L. Hammen and K. M. Yorkston, "Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria," *Journal of communication disorders*, vol. 29, no. 6, pp. 429–445, 1996.

[16] F. S. Juste, F. C. Sassi, J. B. Costa, and C. R. F. de Andrade, "Frequency of speech disruptions in Parkinson's disease and developmental stuttering: A comparison among speech tasks," *Plos one*, vol. 13, no. 6, p. e0199054, 2018.

[17] S. G. Adams, A. Dykstra, M. Jenkins, and M. Jog, "Speech-to-noise levels and conversational intelligibility in hypophonia and Parkinson's disease," *Journal of medical speech-language pathology*, vol. 16, no. 4, pp. 165–173, 2008.

[18] K. Tjaden and G. Wilding, "Effects of speaking task on intelligibility in Parkinson's disease," *Clinical Linguistics & Phonetics*, vol. 25, no. 2, pp. 155–168, 2011.

[19] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of speech and hearing research*, vol. 12, no. 2, pp. 246–269, 1969.

[20] J. Hlavnicka, R. Cmejla, T. Tykalova, K. Sonka, E. Ruzicka, and J. Rusz, "Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder," *Scientific reports*, vol. 7, no. 1, p. 12, 2017.

[21] E. M. Guilherme, R. d. F. C. Moreira, A. de Oliveira, A. M. Ferro, V. A. P. Di Lorenzo, and A. C. L. Gianlorenco, "Respiratory disorders in Parkinson's disease," *Journal of Parkinson's Disease*, vol. 11, no. 3, pp. 993–1010, 2021.

[22] J. H. Friedman, "James Parkinson did NOT describe respiratory changes in the shaking palsy," *Parkinsonism & Related Disorders*, 2023.

[23] J. H. Won, S. J. Byun, B.-M. Oh, S. J. Park, and H. G. Seo, "Risk and mortality of aspiration pneumonia in Parkinson's disease: A nationwide database study," *Scientific reports*, vol. 11, no. 1, p. 6597, 2021.

[24] A. Docu Axelerad, A. Z. Stroe, O. C. Arghir, D. Docu Axelerad, and A. E. Gogu, "Respiratory dysfunctions in Parkinson's disease patients," *Brain Sciences*, vol. 11, no. 5, p. 595, 2021.

[25] K. Kaczyńska, M. E. Orłowska, and K. Andrzejewski, "Respiratory abnormalities in Parkinson's disease: What do we know from studies in humans and animal models?" *International Journal of Molecular Sciences*, vol. 23, no. 7, p. 3499, 2022.

[26] A. D'Arrigo, S. Floro, F. Bartesaghi, C. Casellato, G. F. S. Papa, S. Centanni, A. Priori, and T. Bocci, "Respiratory dysfunction in Parkinson's disease: a narrative review," *ERJ Open Research*, vol. 6, no. 4, 2020.

[27] Z. Galaz, J. Mekyska, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, and D. Berankova, "Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease," *Computer methods and programs in biomedicine*, vol. 127, pp. 301–317, 2016.

[28] M. Kostalova, M. Mrackova, R. Marecek, D. Berankova, I. Eliasova, E. Janousova, J. Roubickova, J. Bednarik, and I. Rektorova, "The 3f test dysarthric profile-normative speach values in czech," *CESKA A SLOVENSKA NEUROLOGIE A NEUROCHIRURGIE*, vol. 76, no. 5, pp. 614–618, 2013.

[29] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]. version 5.3. 51," *Online: http://www. praat. org/retrieved, last viewed on*, vol. 12, 2013.

On Hyperspectral Analysis of Water Soluble Writing Inks

1st Tomáš Zemčík
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
zemcikt@vutbr.cz

2nd Karel Horák
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
horak@vutbr.cz

***Abstract*—In this short paper a possibility of using Hyperspectral Imaging based analysis of water soluble inks is investigated. The current research has investigated this method mostly for marker inks, ballpoint pen inks and gel inks, yet curiously the whole water soluble ink class seems underrepresented. Samples were prepared and hyperspectral images of said samples taken, and some preliminary statistical analysis was performed. The data at this stage suggests that at least some samples are visually distinguishable from other samples, while certain other samples are virtually indistinguishable. A method for applying this knowledge in a handwritten text ink identification is briefly outlined and implemented in a simplified way, the results are promising on good input data, but poor on subpar images. The weaknesses of the method are discussed and potential improvement strategies proposed.**

***Index Terms*—hyperspectral imaging, specim IQ, ink analysis, document authentication**

I. Introduction

In the field of document analysis differentiation of various types of ink is an important part of understanding the way said document was created, edited, or perhaps further manipulated over it's lifespan. This understanding is important for a wide field of areas, such as historical document analysis, forensics or restoration [1].

There are numerous established methods for analysing inks. Menżyk et al. in [2] give an exhaustive overview of ink analysis methods in forensic examination of documents, and divide the methods into following categories in order of inquiry:

- **Preliminary examinations** - non-destructive methods requiring little to no specialised equipment. These include visual examination using microscopes or stereomicroscopes, optical analysis using various light sources including in the UV spectra.
- **Chromatographic methods** - methods of chemical analysis that require invasive sample collection from the document, and its further processing by dissolving the ink in appropriate solvent and performing the chromatography. The most common methods are Thin Layer Chromatography (TLC), High-Performance Thin Layer Chromatography (HPTLC), High-Performance Liquid Chromatography (HPLC) and Gas Chromatography (GC).
- **Mass spectrometry** ([2] lists mass spectrometry as part of spectroscopic methods, but for the different characteristics this paper distinguishes it from other spectroscopic methods) - destructive analysis of a sample to determine its precise chemical composition.
- **Spectroscopic methods** - the latest group of methods analysing spectral properties of inks. Spectroscopic methods used in document analysis include Fourier Transform Infrared Spectroscopy (FTIR), Raman spectroscopy and UV-VIS spectroscopy. Hyperspectral Imaging (HSI) is another representative of spectroscopic methods, which on top of the spectral information provides the spatial distribution of the information.

In the current ink analysis state-of-the-art as outlined by [2], and considering the recent developments in HSI techniques, the attention has been mostly concentrated on ballpoint inks, gel inks, printing inks and other kinds of artwork [3], [4]. Water soluble inks, such as drawing inks, calligraphy inks and fountain pen inks have gone mostly unexamined by the HSI methods. In this paper the results of the measurement are presented, and a proof-of-concept application is outlined and discussed to evaluate the feasibility of HSI based analysis on some selected water soluble inks.

II. Materials and methods

A. The sample preparation

Sample cards of the 10 analysed document inks were prepared. The inks were selected to represent some of the most common shades of blue document ink. Deliberately such shades were selected that in some cases they cannot be distinguished by the naked eye, and that depending on the application one could not even discern which of the ink is darker or lighter regardless of the shade. In addition a number of sample cards with marker inks was prepared to compare against the document inks as a reference and control group.

The individual samples were fashioned on a standard 80gsm printing paper representing the most common and available substrate. A square of ink was applied with a brush to supply a sufficient surface for individual pixel spectra analysis. The brush application was selected deliberately despite the fact that it is a rare way to apply ink in practice, it was selected to create areas of varying thickness of application from heavier to lighter, no attempt was made to create a gradient or any organised sort of pattern, but the nature of the brush application intrinsically results in the desired pattern. In addition a sample text applied with a fine nib is included, and a small square of ink applied very thickly with a nib is also included as a high density application example. All sample coupons were allowed to dry for a period of 24 hours before any measurement. The marker samples were prepared to the identical design, excepting the marker was used, and therefore various patterns in the sample square, and thickness in the sample text occur.

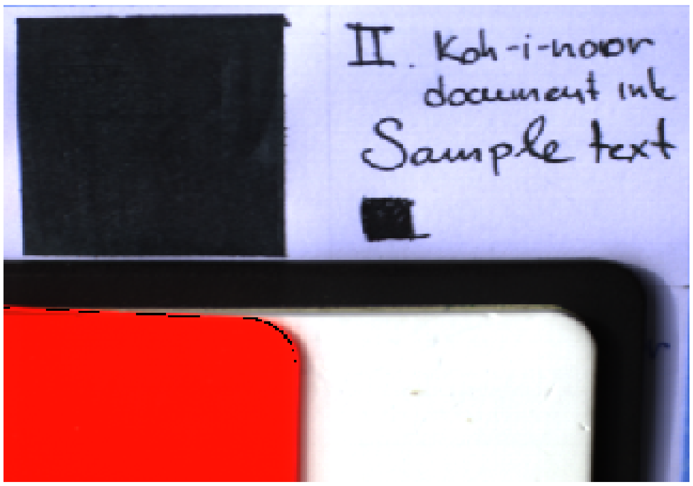


Fig. 1. Sample as prepared and captured (in pseudo RGB colours and cropped)

B. Image acquisition

The described samples were imaged using the Specim IQ hyperspectral camera. The camera is one of the first portable hyperspectral imaging solutions, and was released in 2018 by the Oulu, Finland based Specim company [5]. A multitude of applications have already utilized this camera, such as [6]–[9].

TABLE I
SPECIM IQ BASIC SPECIFICATIONS

Specim IQ basic specifications	
Parameter	**Value**
Spatial resolution	512 x 512 px
Spectral resolution	7 nm
Spectral range	400 - 1000 nm
F/number	1.7
Data format	ENVI compatible
Battery	5200 mAh Li-ion
Battery life	100 captures
Focusing distance	150+ mm
FoV	31° x 31°
Operating temperature	0°C - +40°C

The samples were placed slightly beyond the minimum focus distance of 200 mm from the lens, and illuminated sequentially with sunlight and artificial lighting using a special spectrum-controlled illumination unit tuned to as uniform a spectrum as possible. Calibration reference was included in the scene consisting of a white chip with known uniform reflectance across the spectral range of the camera, and a orange chip of a known optical properties with the reflectance cutoff of 590nm.

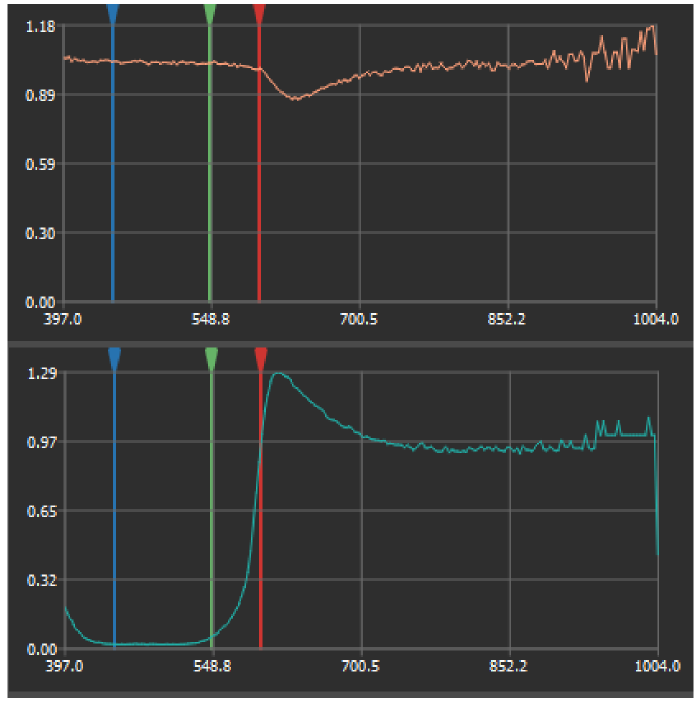


Fig. 2. The measured spectra of the reference chips, the white reference chip on top, the orange reference chip on the bottom

The measured spectra were verified using the OceanOptics USB 2000+ fibre optics spectrometer using the reflective spectrometry method.

The measurements were repeated for all 10 writing ink samples, and 4 control marker ink samples. Captures with multiple samples in a single image were also taken in order to allow for later cross-check between the various samples for variation of illumination (in case of the daylight samples, since the captures and sample changes take a non-negligible amout of time and the lighting conditions may have shifted in the meanwhile).

III. DATA ANALYSIS

The individual scanned samples were converted to a hypercube image representation of the dimensions 512 x 512 x 204 - representing the 204 spectral slices of 512 px x 512 px each. Within each sample picture the sample ink square of uniform dimensions of 150 px x 150 px was annotated by hand, and a subimage hypercube of the dimensions 150 x 150 x 204 was extracted as a data sample. This represents 22500 individual pixels each containing a full spectrum, where each pixel itself is a sample of that particular place's application of ink. Together they represent a representative sample of ink for further statistical analysis. Such samples may be collected in a database of inks and be used for classification of unknown samples.

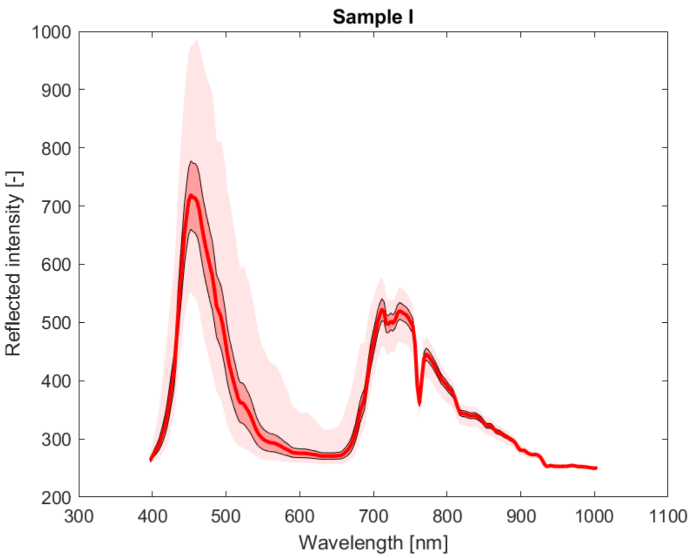


Fig. 3. Sample I

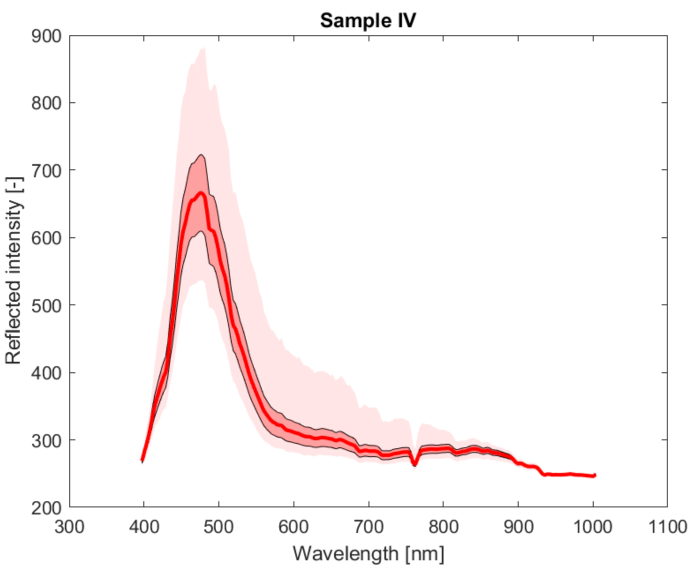


Fig. 4. Sample IV

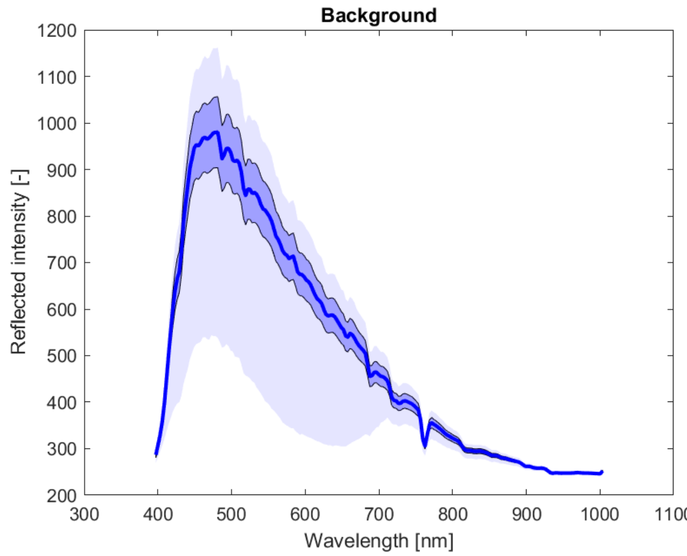


Fig. 5. Background sample

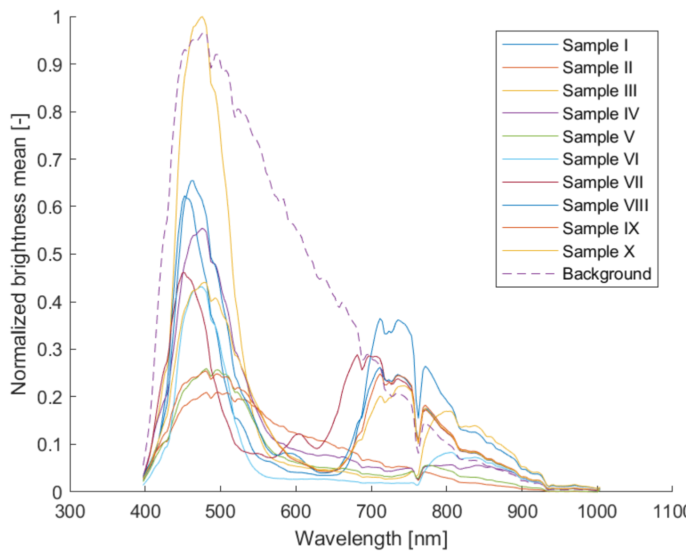


Fig. 6. Sample II

In figures 3 through 5 two ink samples (in red) and a statistical sample of the background (in blue) are represented. The solid line represents the mean value of intensity across the 22500 pixels (a variable higher number of pixels is used for the background sample) for each of 204 spectral bins between circa 400 nm to 1000 nm. The darker shaded area around the mean line represents the standard deviation area across the samples, and the lighter shaded area represents the absolute minimum and maximum for each sample.

In figure 6 the normalised mean spectra of all ten ink samples (in solid lines), and the background (in dashed line) are given. Please note that some of the mean spectra extend above the background, which in reflective spectroscopy would imply the unlikely situation of the inked surface having higher reflectivity than the (white) background. This arises from the slight variance of ambient illumination during image acquisition and from the nature of the statistical analysis used (the background sample was large and included more or less shadowed areas).

Even by visual inspection it is clear that some of the selected samples and the background (as shown above) are easily distinguishable from certain other samples, therefore there is clearly a potential for discrimination between at least some combinations of samples.

IV. The proposed method

The proposed proof-of-concept method is intended to be a simple way to demonstrate that a solution of water soluble ink differentiation exists within the domain of HSI.

For a future usable application, the presumed form of the analysed sample would be a block of handwritten text, or a drawing on some sort of a substrate. During the sample imaging it would be important to ensure a correct acquisition of both the ink application areas, and a sufficient sample of the background for future decomposition. In the current simplified experiment an identical background (a standard copier paper) is assumed.

In figure 7 a workflow of the proposed simple method is outlined, the individual blocks are then described in the paragraphs below:

After **image acquisition** the sample file is collected from the camera and converted into a hypercube representation. Multiple step **preprocessing** is then performed on the image. First a Region of Interest (RoI) is selected, and the hypercube cropped to dimension in order to reduce computational requirements. Then Gaussian blur of appropriate kernel size (depending on the definition of the finest lines) is applied over all the spectral images. This is done in order to avoid having certain spectral slices in good focus, while other slices are our of focus due to the nature of different wavelengths propagating through materials at different velocities, and the imperfect compensation of the quartz lens used. Here a additional performance may be gained by analysing which wavelengths suffer what level of blur and then blurring the other wavelengths accordingly, the distribution will not change significantly with camera settings. And finally known bands of interest may be removed or selected based on a priori knowledge of the materials.

Then **background selection** is performed manually on the indicated image, an area as large as possible of the clear substrate is indicated to the algorithm for a statistical sample to be collected.

Optionally a **dimensionality reduction** step may be included to further reduce the hypercube size. On the analysed samples of water soluble inks 10 principal components have proved to carry well in excess of 90% of the information, therefore this step could offer a significant increase in the system's performance.

Using the previously collected statistical samples of the inks (as described above) and a selected background sample a **Support Vector Machine (SVM) classifier** model is trained to classify individual pixels in the image. Applying this model pixelwise a label image is created with individual pixels representing a class prediction for the corresponding hypercube spectrum.

Because the pixelwise classification is is not perfect on real data, a step of **smoothing and filtering** the resulting label image is taken. By applying a sequence of elementary morphological operations extremely small areas of discrete classes are removed (usually the borderline pixels mixing background and ink spectra or spectra of different inks). Larger areas of single labels are thus smoothed and merged where necessary. The output label image is then produced.

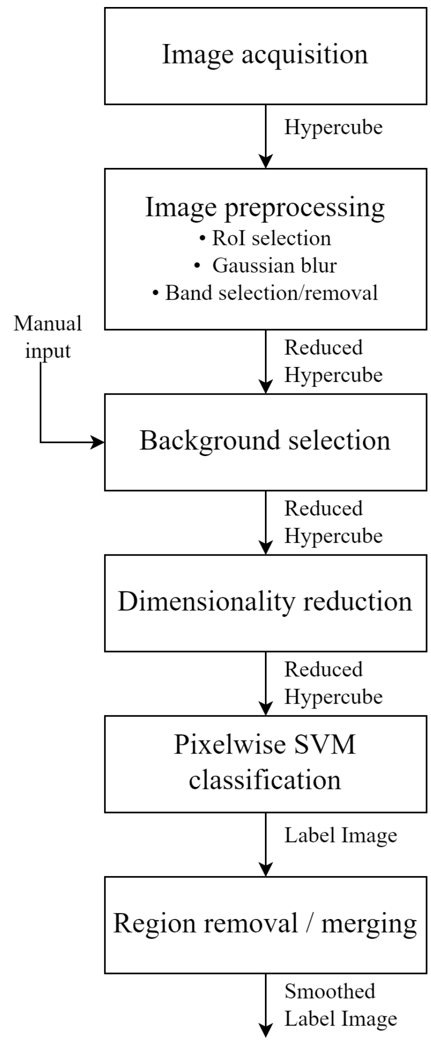


Fig. 7. The simplified proposed method workflow

A. Results and discussion

In figure 8 a successfully produced unsmoothed label image is shown. Note that while Sample III is near perfectly classified, segmented and resolved, the Sample V below is near perfect in the large brush applied area and pen applied square, however the pen applied sample text and description are not well resolved, and missclassified pixels (confused with Sample II in light blue, and Sample IV in red) are also observed. Generally the methods seems to perform fairly well on large resolution images such as this one, some combinations of inks are prone to miss-classification more than others, as expected (subject to future quantitative analysis).

Figure 9 demonstrates a loss of resolving power when an insufficient resolution and lighting image is supplied. The sample VI in blue is clearly being confused with sample III in purple, and Sample IV in red in case of written text. The background spectra, while still correctly classified, have been

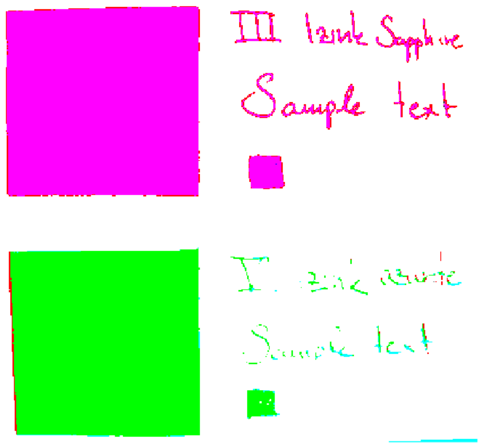


Fig. 8. Unsmoothed label image as a result on good resolution source data (Sample III in purple, Sample V in green and the Background in white)

variable throughout the region of interest due to the insufficient lighting.

Fig. 9. Unsmoothed label image as a result on insuffucent resolution source data (Sample IV in red, Sample VI in blue and the Background in white)

The method seems to perform fairly poorly when either the resolution of handwritten text, or the lighting conditions are compromised. To remedy the lighting variability an incorporation of the calibration procedure outlined above using the calibration object included in the samples is necessary, experimentation with calibrated images showed improvement in classification quality, however the results have not yet been stable.

The poorer resolution caused problems may to an extent be mitigated by employing hyperspectral endmember extraction methods as outlined in [10]. These methods are designed to decompose spectra of individual pixels into different materials (endmembers) in the image, this helps in resolving the fringe regions where in a single pixels spectra of multiple materials are projected (due to insufficient resolution, blur or other factors). This would be particularly helpful for the thin lined handwritten text.

V. Conclusions

In this short paper a method for preparing water soluble writing ink samples, and capturing their hyperspectral representation using the Specim IQ camera is presented. Examples of the results obtained are shown, including some basic statistical analysis, and a method for automated ink sample recognition is outlined.

Based on the experiments performed the author believes that discrimination between different water soluble inks using a hyperspectral imaging method is (at least in some cases) feasible. In some cases the spectra are virtually identical, and a deeper analysis is required to determine whether or not a solution within the hyperspectral imaging domain (or within the Specim IQ camera spectral range) is possible. Also the selected pixel-wise approach is not perfect for discriminating the ink type for given object in an image, a more sophisticated segmentation and classification approach would be required in the future application.

A method for analysing samples of handwritten text is briefly outlined and discussed. This proof-of-concept method, while being very far from being a usable application, shows it is certainly possible to differentiate certain water soluble inks by way of hyperspectral image analysis - as speculated above. In future work the author intends to incorporate the improvements outlined in the previous chapter (input image calibration and endmember extraction methods), and expand the experimentation on larger sample of handwritten text using more inks in order to perform a quantitative analysis of differentiation ability of this method, it would also be interesting to expand the experimentation into other colours of ink, such as blacks and reds.

Acknowledgment

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

References

[1] Morales, Aythami & Ferrer, Miguel & Diaz, Moises & Carmona-Duarte, Cristina & Thomas, Gordon. (2014). “The use of Hyperspectral Analysis for Ink Identification in Handwritten Documents“. Proceedings - International Carnahan Conference on Security Technology. 2014. 10.1109/CCST.2014.6986980.

[2] Menżyk, Alicja & Zadora, Grzegorz & Sajewicz, M.. (2016). “Physicochemical analysis of inks and methods of evidence evaluation - A review“. Problems of Forensic Sciences. 104. 245-278.

[3] Tzeng, Di-yuan & Berns, Roy. (2001). “Spectral-Based Ink Selection for Multiple-Ink Printing I. Colorant Estimation of Original Objects“.

[4] Maria Kubik, Chapter 5 “Hyperspectral Imaging: A New Technique for the Non-Invasive Study of Artworks“, Editor(s): Dudley Creagh, David Bradley, Physical Techniques in the Study of Art, Archaeology and Cultural Heritage, Elsevier, Volume 2, 2007, Pages 199-259, ISSN 1871-1731, ISBN 9780444528568, https://doi.org/10.1016/S1871-1731(07)80007-8.

[5] E. Ikäheimo and J. Jussila, ”Introducing Specim IQ”. Oulu, FInland: Specim, Spectral Imaging Oy, 2018.

[6] J. Behmann et al., “Specim IQ: Evaluation of a New, Miniaturized Handheld Hyperspectral Camera and Its Application for Plant Phenotyping and Disease Detection,” Sensors, vol. 18, no. 2, p. 441, Feb. 2018, doi: 10.3390/s18020441.

[7] Solovchenko, Alexei & Dorokhov, Alexei & Shurygin, Boris & Nikolenko, Alex & Velichko, Vitaly & Smirnov, Igor & Khort, Dmitry & Aksenov, Aleksandr & Kuzin, Andrei. (2021). ”Linking Tissue Damage to Hyperspectral Reflectance for Non-Invasive Monitoring of Apple Fruit in Orchards”. Plants. 10. 310. 10.3390/plants10020310.

[8] Tejada-Casado, Maria & Ghinea, Razvan Ionut & Martínez-Domingo, Miguel & Pérez Gómez, María del Mar & Cardona, Juan & Ruiz-López, Javier & Herrera, Luis. (2022). ”Validation of a Hyperspectral Imaging System for Color Measurement of In-Vivo Dental Structures”. Micromachines. 13. 1929. 10.3390/mi13111929.

[9] Alt, V. & Gurova, Tamara & Elkin, O. & Klimenko, D. & Maximov, L. & Pestunov, I. & Dubrovskaya, O. & M.A., Genaev & Erst, T. & Genaev, K. & Komyshev, Evgenii & Khlestkin, Vadim & Afonnikov, Dmitry. (2020). ”The use of Specim IQ, a hyperspectral camera, for plant analysis”. Vavilov Journal of Genetics and Breeding. 24. 259-266. 10.18699/VJ19.587.

[10] K. V. Kale, M. M. Solankar, and D. B. Nalawade, “Hyperspectral Endmember Extraction Techniques,” Processing and Analysis of Hyperspectral Data, Jan. 2020, doi: 10.5772/intechopen.88910.

Field emission characteristics and analysis of charge flow through graphite based cathodes

1st Mohammad M. Allaham
Central European Institute of Technology
Brno University of Technology
Institute of Scientific Instruments of Czech Academy of Sciences
Brno, Czech Republic
allaham@vutbr.cz

2nd Daniel Burda
Department of Physics
Faculty of Electrical Engineering and Communication
Institute of Scientific Instruments of Czech Academy of Sciences
Brno, Czech Republic
burda@isibrno.cz

3rd Dinara Sobola
Central European Institute of Technology
Brno University of Technology
Department of Physics
Faculty of Electrical Engineering
Brno, Czech Republic Sobola@vut.cz

4th Alexandr Knápek
Institute of Scientific Instruments of Czech Academy of Sciences
Brno, Czech Republic
knapek@isibrno.cz

***Abstract*—This paper studies the performance of different types of graphite cathodes when operated as field emission electron sources. The tested cathodes were prepared in the form of bulk polymer graphite, bulk pure graphite, and polymer graphite coated with thin layer of insulating material (epoxy resin). The obtained results include X-Ray photoelectron spectroscopy analysis, scanning electron micrographs, the field emission microscope patterns and current-voltage characteristics, and the orthodoxy test analysis results. The importance of this study is summarized in following the pursuit of finding cheap and green electron sources. Moreover, to present a new exotic field emission behavior in the form of emission pulses.**

***Index Terms*—field emission, pulsed field emission, graphite, microflakes, polymer graphite.**

I. Introduction

Field electron emission phenomenon is a special technique in microelectronics. This phenomenon appears when the electric current flows in vacuum gap between two separated electrodes after applying sufficient voltage. The external electric field in the gap reduces the potential energy barrier (PEB) near the surface of the two electrodes allowing electrons to quantum tunnel through at sufficient voltage values. Fig. 1 shows the changes in the shape of the PEB from rectangular to image rounded (IR-PEB) after applying the Schottky image effect, and finally to the reduced image rounded or Schottky-Nordheim potential energy barrier (SN-PEB) after applying an external electric field [1].

In field emission studies, there are two types of cathodes (electron sources): The single tip field emitter (STFE) and the large area field emitters (LAFE), which are synthesized by an array of micro/nano STFEs. In what follows, the model to describe the emission current from LAFE is used since the graphite samples are composed of a large array of microflakes with several pointed nano-tips [2].

The height of the SN-PEB in Fig. 1 is described by the electron motive energy $M^{\mathrm{SN}}(z)$. After solving the Schrodinger equation for the presented system, the measured emission current (I) is described as a function of the applied measured voltage (V) and the local work function (ϕ) of the used material using Forbes extended Murphy-Good equation [3]:

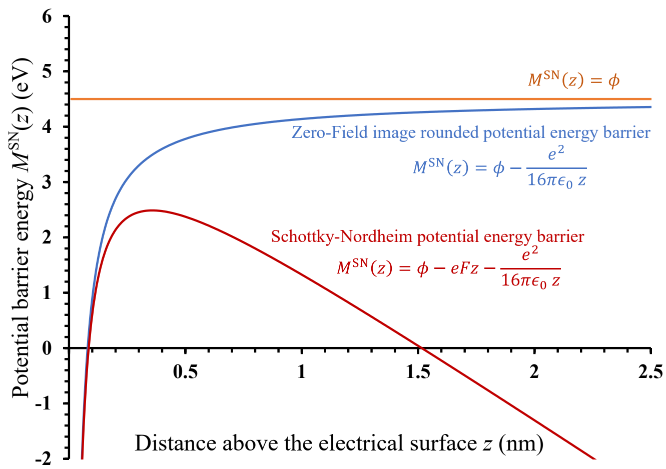


Fig. 1. Schematic for the triangular, image-rounded zero-field barrier and the Schottky-Nordheim (SN) barrier. The Schottky-Nordheim barrier is evaluated for ϕ= 4.50 eV and at $F \approx 2.8$ V/nm.

$$I(V) = A_{\mathrm{f}}^{\mathrm{SN}} a \phi^{-1} \zeta^{-2} V^2 \exp\left(\frac{-\mathrm{v_F} b \phi^{3/2} \zeta}{V}\right) \tag{1}$$

In eq. 1, $A_{\mathrm{f}}^{\mathrm{SN}}$ is the formal emission area from a SN-PEB, ζ is the voltage conversion length that connects between the applied electric field and the measured voltage, a and b are the first and second Fowler-Nordheim constants, and $\mathrm{v_F}$ is a special mathematical function known as the shape correction factor [4]. Using eq. 1, Fig. 2 presents the distribution of the measured current as a function of the measured voltage

DOI 10.13164/eeict.2023.243

and the local work function. In 2019, Richard G. Forbes introduced a modern and more accurate analysis and testing method known as the Murphy-Good analysis plots [1], [4]. This analysis method provides a nearly exact straight line, which provides more accurate extraction procedure for the characterization parameters $A_{\mathrm{f}}^{\mathrm{SN}}$ and ζ. MG-plots have the form $\ln(I.V^{-\kappa})$ vs V^{-1} with κ as the pre-exponential factor of the MG-plot as can be evaluated if ϕ is known using the scaling parameters calculator in [5]. From a MG-plot, ζ can be obtained from the slope and $A_{\mathrm{f}}^{\mathrm{SN}}$ can be obtained from the vertical axis intercept value [6].

Graphite is an adjacent multilayer of graphene, which is synthesized by parallel planar network of carbon atoms ordered in an uncentered hexagonal structure. The carbon atoms are covalently bounded via the van der Waals forces [7]. Polymer graphite is a mixture of graphite microflakes with a polymer bonding agent, which in general is white clay or (Kaolinite) $\mathrm{Al_2Si_2O_5(OH)_4}$ [8].

In this paper, the field emission characteristics from polymer graphite, pure graphite, and coated polymer graphite will be presented. This includes the current-voltage characteristics $I(V)$ and the MG analysis results. Scanning electron micrographs (SEM) are presented to show the difference in structure between polymer and pure graphite samples. Moreover, the SEM results are supported by X-Ray photoelectron spectroscopy (XPS) analysis to prove the purity of the used pure graphite microflakes.

II. Methodology

A. Materials

1) Pure graphite microflakes cathodes: Graphite microflakes (GMF) are formed by the incomplete combustion (limited supply of Oxygen) of hydrocarbon gases and vapors derived from petroleum sources at temperature of nearly 1400 $^\circ\mathrm{C}$. In this work, the used GMF has the commercial name Vulcan® XC-72 produced by Cabot Inc. The product is mainly high purity 99.99% graphite nanoparticles that can be achieved after ultrasonic treatment of the GMFs.

To prepare the pure graphite cathodes, micro-glass tubes were prepared using the heating-pulling technique. Glass tubes of inner diameter $\sim$1 mm were attached to a holder from one side, and to a pulling drill chuck from the other side. Electrical current is then applied into the heating loop to melt the glass while being pulled down at the same time by the chuck. This process can help to pull the melted glass creating micro-conic tips in the form of micro-glass tubes with inner diameters of approximately $40\,\mu\mathrm{m}$.

The GMFs were then added inside the glass tube, and ultrasonic treatment was applied to the prepared samples to ensure the GMFs will reach micro-aperture of glass tube. Finally, stainless steel tube of 0.3 mm in diameter was mounted inside the glass tube from the other side to establish the electrical contact between the GMFs and the sample holder (power feedthrough). The results are presented in Fig. 3.

2) Polymer graphite: Polymer graphite cathodes (PGC) can be obtained from the market as micro pencil leads. The combination of polymer and GMFs makes it hard enough to be possible to use. In this paper, authors prepared the required PGCs using the membrane-electrochemical etching technique as described in [2], [9]–[11].

Fig. 4 presents the SEMs of GMF and PGC samples. The structure of GMF is presented in Fig. 4(a and b), where it shows the amorphous structure of pure microflakes. In comparison to the structure of GMF, Fig. 4(c and d) present the structure of PGC, where the existence of the polymer material is clearly visible. In both cases, it is clear that micro-surface of the cathode tip contains large number of random tips act as emission cites, which is a reasonable reason to consider each tip as LAFE.

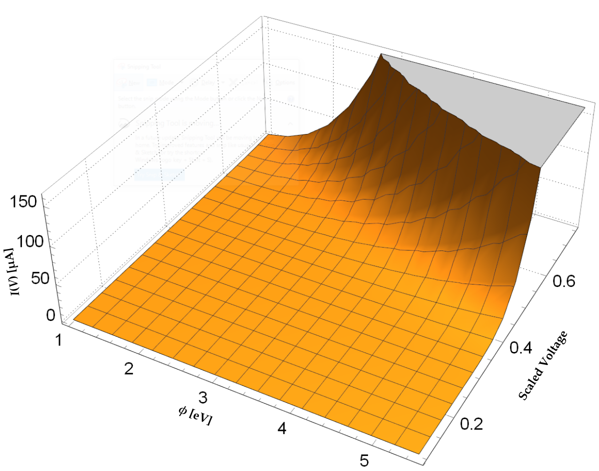


Fig. 2. The distribution of the emission current as a function of the measured voltage and the local work function.

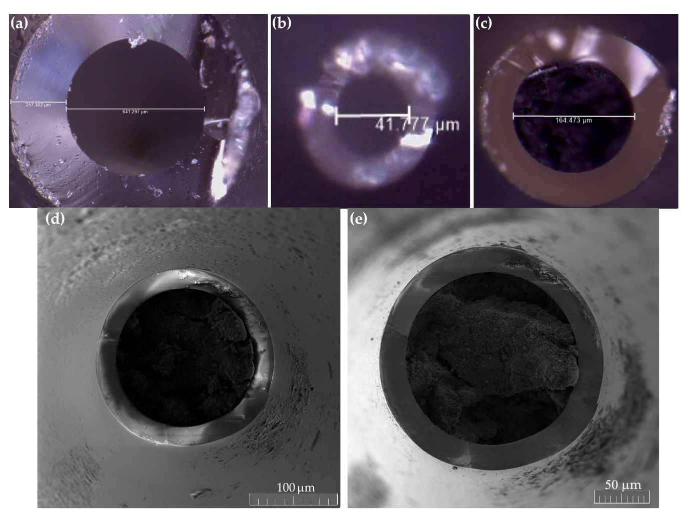


Fig. 3. The micro-glass tubes, where (a) base glass tube, (b) empty micro-glass tip, (c-e) filled micro-glass tube after ultrasonic treatment.

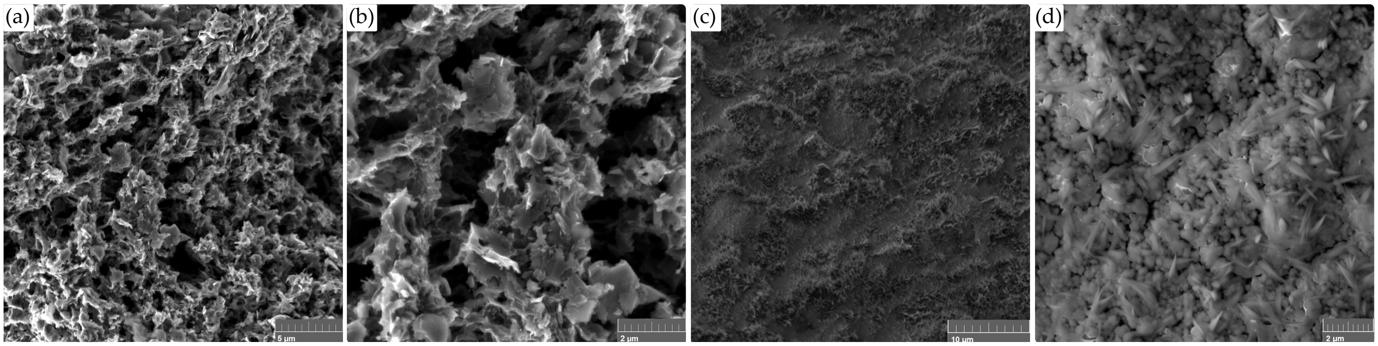


Fig. 4. Scanning electron micrographs of high purity polymer graphite microflakes in comparison to the surface of the polymer graphite pencil lead. (a) shows the random distribution of the microflakes, (b) magnified view of the microflakes, (c) the surface of polymer graphite pencil lead, and (d) magnified view of the microflakes surrounded by the polymer nanoparticles.

XPS analysis were obtained using the AXIS SupraTM X-ray photoelectron spectrometer set-up (KRATOS Analytical Ltd., Manchester, UK). The results are presented in Fig.5(a) for the GMF and Fig.5(b) for the PGCs. For the case of GMFs, the results show only tiny amount of oxygen ($\approx 4\%$) and major percentage of pure carbon ($\approx 96\%$). On the other hand, the results for PGCs show higher oxygen percentage ($\approx 13\%$) with additional tiny percentage of Si ($\approx 2.3\%$) and Al ($\approx 1.3\%$), which came from the the Kaolinite.

3) Coated polymer graphite: To prepare the coated PGCs, Elantas PDG Epoxylite® E478 Single-Component Thixotropic Epoxy VPI Resin is used as the coating thin insulation layer. The epoxy is manufactured for high electrical insulation applications with voltages less than 7000 V. The prepared PGCs were slowly dipped inside the epoxy resin to create a thin layer of the epoxy, which will differ in thicknesses due to the random distribution of the microflakes. The coated samples are then baked at 423 K temperature for 6 hours to cure the epoxy layer.

Because of the variation in the thickness of the coating layer from point to point on the cathode tip surface, the ability for electrons to tunnel through the layer will also vary, making it possible for different tunneling phenomena to occur causing different emission mechanisms, which will be discussed later in the following section.

B. Experimental setup

The prepared samples were mounted as cathodes into a traditional and diode configuration field emission microscopes (FEM) [9]. During the experiments, the pressure inside the FEM was held at ultra-high vacuum conditions (10^{-7} Pa). The anode of the system was used as emitted electrons collector and prepared by coating a YAG scintillator with aluminum. The anode was connected to a grounded auto-ranging picoampermeter (Keithely 405 autoranging Picoammeter) to measure the current-voltage characteristics. Moreover, an autoranging picoammeter branded by Rbd 9103 was also used to record the current-time characteristics at fixed voltage for the case of the coated PGCs.

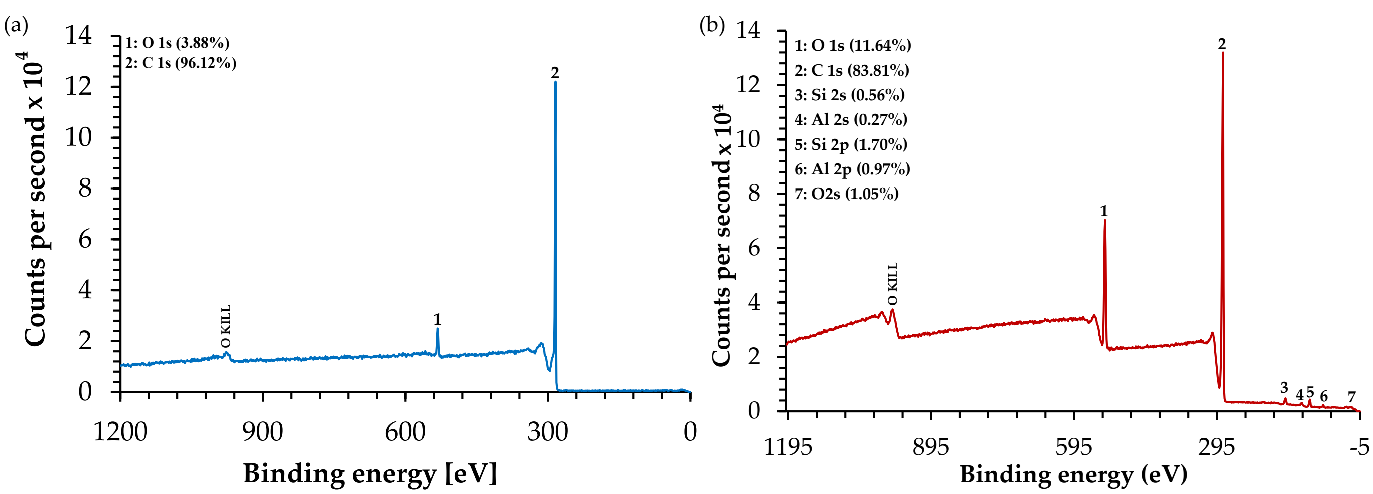


Fig. 5. X-ray photoelectron spectroscopy of (a) 99.99% high purity polymer graphite microflakes, and (b) polymer graphite pencil lead.

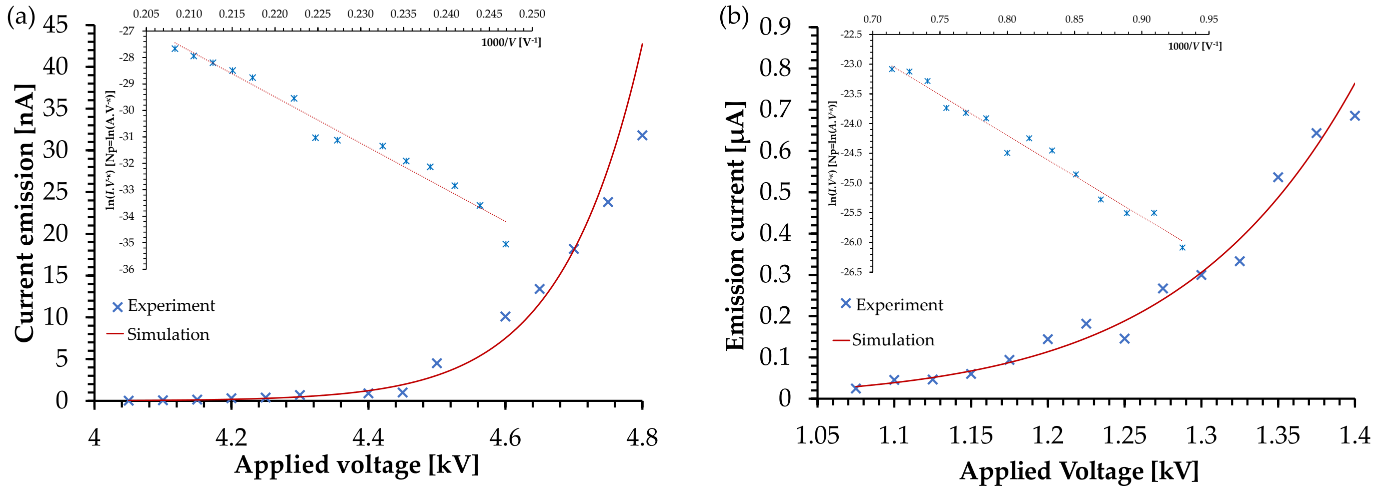


Fig. 6. The current-voltage characteristics and the corresponding analysis plots (insets of each part) for (a) pure graphite microflakes cathodes, and (b) polymer graphite cathode.

III. Results and Discussion

The characteristics of each type of the cathodes were discussed earlier in section II-A of this paper. The results included SEM and XPS analyses for both types of the samples.

In this section, the current-voltage ($I(V)$) characteristics were obtained and presented for the GMF and PGC cathodes using the FEM. The analysis results were obtained using the MG-analysis plots by considering $\phi \approx 4.5$ eV as presented in literature for graphite based materials [2]. The FEM results inluded the field emission patterns that describe the emission current density distribution of the emitted electrons.

For the case of the coated PGC, the current-time ($I(t)$) characteristics were obtained and presented instead of the traditional $I(V)$. This is because the emission process showed a unique behavior as presented later. The applied voltage was fixed at specific voltages where the emission is stable, and the $I(t)$ were recorded.

A. Current-Voltage characteristics

The obtained $I(V)$ for the GMF sample is presented in Fig. 6(a) and for the PGC it is presented in Fig. 6(b). The inset of each figure describes the MG-plot analysis plots. For the presented graphs, it is clear that the PGC has more advantageous results since it has much lower threshold voltage (1075 V vs 4000 V) and much higher emission currents at lower voltages ($\approx$ 900 nA at 1400 V in comparison to 35 nA at 4800 V). To perform the analysis procedure for LAFEs, the macroscopic area (A_{M}) for each tip surface were evaluated and it was 2.1×10^{-8} m^2 for the GMF samples, and 1.6×10^{-9} m^2 for the PGC sample. In both cases the macroscopic distance (d_{M}) between the two electrodes of the FEM was 1 mm. The analyses results were obtained using the field emission analysis webtool in [5], and the results are presented in Table I. The results include the field enhancement factor ($\gamma = d_{\mathrm{M}}/\zeta$) and the formal area efficiency ($\alpha_{\mathrm{f}}^{\mathrm{SN}} = A_{\mathrm{f}}^{\mathrm{SN}}/A_{\mathrm{M}}$), which are

TABLE I
MG-analysis results as obtained for the graphite microflakes (GMF) and polymer graphite (PGC) cathodes.

sample	Slope Np.V	ζ nm	γ	$A_{\mathrm{f}}^{\mathrm{SN}}$ m^2	$\alpha_{\mathrm{f}}^{\mathrm{SN}}$
GMF	-1.7×10^{5}	2600	385.3	1.2×10^{-7}	5.9
PGC	-1.4×10^{4}	213	4688.9	5.1×10^{-18}	3.2×10^{-9}

important parameters to characterize the the used system in terms of enhancing the electric field near the tip surface and the contribution of the tip surface area in the emission process. From Table I, in the case of PGC, the value of $\alpha_{\mathrm{f}}^{\mathrm{SN}}$ lies within a reasonable theoretical range. However, in the case of GMF, the formal emission area has larger values, which indicates that the contribution of electron emission was much higher than the case of PGC. This is also visible from the value of $\alpha_{\mathrm{f}}^{\mathrm{SN}}$ for this sample since it has very high value. This result is important to explain the reason behind achieving lower current values at higher voltages for the case of GMF, since the charge density is distributed among a higher surface area, and thus, electron occupation of the energy states were lower and needed higher supply voltages to reduce the SN-PEB. Moreover, γ value for the case of GMF is $\approx$ 10 times lower than the case of PGC, which means higher supply voltages were needed to provide sufficient field intensities.

B. Current-time characteristics

Following the results presented in section III-A, the PGC were used as base material prepare PGC-epoxy composite cathodes, since PGC provided higher current densities at lower voltages. For this type of hybrid field emission cathodes, the emission process included the appearance of a quasi harmonic pulses of high current densities. For this reason, the supplied voltages were fixed at 1700 V when the emission current had a

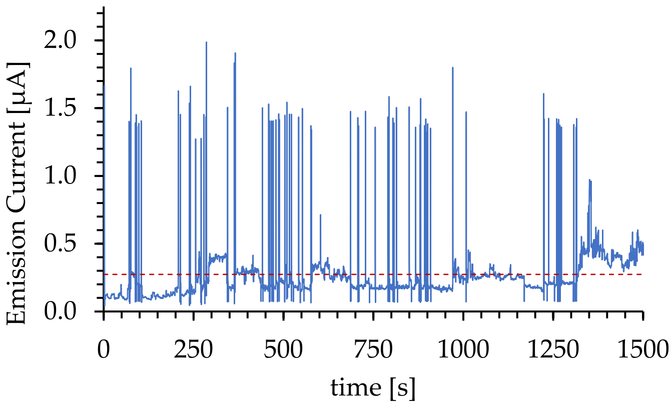


Fig. 7. The quasi-harmonic behavior that was obtained after coating the polymer graphite with thin layer of epoxy resin.

stable value and the $I(t)$ were studied for 1500 s. The obtained results are presented in Fig. 7.

This quasi-harmonic behavior can be explained by considering nano-capacitors that are formed within the coating layer after applying an intense external electric field. The electric field will rearrange the electrical dipoles causing some reorientation process for the direction of the dipoles. As the intensity of the electric field increases, the length of the dipoles will increase which causes to merge the dipoles forming larger dipoles in the form of nano-capacitors. At some point, merging of the formed nano-capacitors will not succeed and instead, a discharge of electrons from the a nano-capacitor to another can occur by a resonate quantum tunneling phenomenon. The latest can occur due to presence of trapped electrons between two potential energy barriers, which can lead to harmonically discharge the electrons from surface of the coating layer to vacuum in the form of intense electron pulses.

IV. Conclusion

In this paper, three types of graphite based field emission electron sources (cathodes) were prepared, tested, and analysed. The prepared cathodes were prepared from pure graphite microflakes, polymer graphite, and polymer graphite coated with epoxy resin. The results showed higher advantageous characteristics for the case of the polymer graphite cathodes and for this reason they were used to produce the coated samples. The coated samples showed a unique field emission behavior by emitting intense pulses of electrons with high current densities.

The poor performance of the pure bulk graphite cathodes is related to the limitation of the current flow inside the bulk structure of graphite microflakes. For this reason, as a future work, the project aims to proceed with pure graphite microflakes and study the field emission characteristics from cathodes prepared from shells of pure graphite microflakes, since the shell structure is believed to provide promising field emission characteristics.

Acknowledgment

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

Research described in the paper was financially supported by the Internal Grant Agency of Brno University of Technology, grants numbers CEITEC VUT/FEKT-J-23-8307 and FEKT-S-20-6352.

The research was financially supported by the Czech Academy of Sciences (RVO:68081731).

References

[1] M. M. Allaham, R. G. Forbes, A. Knápek, D. Sobola, D. Burda, P. Sedlák, M. S. Mousa, Interpretation of field emission current–voltage data: Background theory and detailed simulation testing of a user-friendly webtool, Materials Today Communications 31 (2022) 103654. doi:https://doi.org/10.1016/j.mtcomm.2022.103654.

[2] A. Knápek, R. Dallaev, D. Burda, D. Sobola, M. M. Allaham, M. Horáček, P. Kaspar, M. Matějka, M. S. Mousa, Field emission properties of polymer graphite tips prepared by membrane electrochemical etching, Nanomaterials 10 (7) (2020). doi:10.3390/nano10071294.

[3] R. G. Forbes, Development of a simple quantitative test for lack of field emission orthodoxy, Proc. R. Soc. Lond. A 469 (2013) 20130271. doi:10.1098/rspa.2013.0271.

[4] R. G. Forbes, The Murphy–Good plot: a better method of analysing field emission data, R. Soc. Open Sci. 6 (2019) 190912. doi:10.1098/rsos.190912.

[5] T. Endo, T. Higuchi, Y. Yamada, M. Sasaki, Field emission from mechanical pencil lead and graphite edges, in: IVESC 2012, 2012, pp. 307–308. doi:10.1109/IVESC.2012.6264185.

[6] R. S. Khairnar, C. V. Dharmadhikari, D. S. Joag, Pencil lead tips: A field ion and field electron emission microscopic study, Journal of Applied Physics 65 (12) (1989) 4735–4738. doi:10.1063/1.343225.

[7] F. V. Chukhrov, B. B. Zvyagin, A. P. Zhukhlistov, N. I. Organova, L. P. Yermilova, Structural features of natural graphite, International Geology Review 28 (7) (1986) 790–801. doi:10.1080/00206818609466321.

[8] A. Knápek, D. Sobola, D. Burda, A. Daňhel, M. S. Mousa, V. Kolařík, Polymer graphite pencil lead as a cheap alternative for classic conductive spm probes, Nanomaterials 9 (2019) 1756. doi:10.3390/nano9121756.

[9] A. Knápek, M. M. Allaham, D. Burda, D. Sobola, M. Drozd, M. Horáček, Explanation of the quasi-harmonic field emission behaviour observed on epoxy-coated polymer graphite cathodes, Materials Today Communications 34 (2023) 105270. doi:https://doi.org/10.1016/j.mtcomm.2022.105270.

[10] R. Ławrowski, M. Bachmann, R. Schreiner, Field emission arrays from graphite fabricated by laser micromachining, Journal of Vacuum Science & Technology B 40 (2) (2022) 023203. doi:10.1116/6.0001547.

[11] L. Chen, H. Yu, J. Zhong, L. Song, J. Wu, W. Su, Graphene field emitters: A review of fabrication, characterization and properties, Materials Science and Engineering: B 220 (2017) 44–58. doi:https://doi.org/10.1016/j.mseb.2017.03.007.

Diffusion Behavior of Analyte Molecules in a Nanoporous Matrix Created from Polystyrene Nanoparticles

Zuzana Košelová
Department of Microelectronics
Faculty of Electrical Engineering and Communication,
Brno University of Technology,
Brno, 612 00, Czech Republic
225675@vut.cz

***Abstract*— This study investigates the diffusion behavior of analyte molecules in a nanoporous matrix of polystyrene particles. Two approaches, steady-state current and potential step chronoamperometry were employed to calculate diffusion coefficients and compare the effect of nanoporous matrices with different diameter of pores and different method of immobilization of enzyme. Pores were created with close packing of polystyrene nanoparticles. The covalent and cross-linking immobilization were used to immobilize the enzyme. Chronoamperometric measurement was applied to detect analyte. The study found that the presence of nanoporous matrices immobilized on p-lysine affects diffusion. The potential step chronoamperometry method showed the improvement of diffusion. This study provides insight into the diffusion behavior of analyte molecules in nanoporous matrices, which can be useful in optimizing the design of amperometric biosensors.**

Keywords—Diffuse coefficient, nanoporous matrix, potential step chronoamperometry, polystyrene nanoparticles, steady-state current

I. Introduction

The diffusion coefficient is a fundamental property of a material that characterizes how quickly particles or molecules can move through it under the influence of random thermal motions [1]. It is used to model the transport of solutes in porous media, is important in fields such as environmental science, geology, and biomedical engineering. It can be used to model the transport of analyte molecules to the sensing element of the nanobiosensor. The rate at which the analyte molecules diffuse to the sensing element depends on the diffusion coefficient of the molecules and the properties of the surrounding medium. By understanding the diffusion behavior of analyte molecules, the design of the sensing element can be optimized to improve the sensitivity, accuracy, and speed of the nanobiosensor [2]. In recent years, there has been a growing interest in modifying sensors using nanomaterials, particularly in the field of nanopore membranes for biosensors [3]. Nanoporous materials have unique properties, such as high surface area, tunable pore size, and the ability to selectively capture or transport molecules, that make them attractive for use in biosensors. They have numerous potential applications in areas such as medical diagnostics, environmental monitoring, and food safety. A dense planar arrangement of spherical nanoparticles (NPs) can be a viable option for creating nanopores, as the spaces between the NPs form the pores, with their width determined by the NPs' diameter. This allows to produce a versatile system by selecting the appropriate NPs material and shape. The study presented here utilized nonconductive polystyrene NPs of uniform size to generate porous matrices on electrodes, resulting in a homogenous monolayer. Electrochemical measurements, in the presence of redox mediators, are typically used to quantify the affinity phenomena in these nanopores.

In the past, it was worked with similar systems as this paper [4][5]. We want to test, if diffusion through the pores is an important factor and whether the diffusion coefficients will decrease (a higher diffusion coefficient means better transport of substances). We also wanted to investigate if the sample without NPs, but with enzyme glucose oxidase (GOD) on poly-L-lysine hydrobromide (p-lys) will have the same diffusion. Therefore, we can distinguish only the effect of NPs. Then only the bare electrode with cross-linked enzyme was tested, to determine effect of p-lys separately.

However, it is quite challenging for our electrode arrangement to devise and conduct diffusion coefficient measurements, especially considering the presence of not only ruthenium diffusion as both Hexaammineruthenium(II) chloride (Ru(II)) and Hexaammineruthenium(III) chloride Ru(III) but also glucose and its products. Because due to enzymatic reaction with glucose the ruthenium mediator Ru(III) was reduced to Ru(II) on electrode as showed on Fig. 2A,C. Two approaches for determining diffuse coefficient were selected, both based on amperometry measurements, namely the Steady-state current approach and the Potential step chronoamperometry approach. These are not new approaches, as they have already been employed as models in other systems [2] [6].

II. Theory

A. Steady-state current approach

Steady-state current approach can be used for determining the diffusion coefficient of electroactive species in solution. In this approach, a steady-state current is measured as a function of the concentration of the electroactive species at the electrode surface. The modified expression for the steady-state current was used:

$$I = 4nFcDr_0 \quad (1)$$

And were modified to expression for the microelectrodes with a recess depth [2]:

$$I = \frac{4nFcDr_0}{\frac{4L}{\pi r_0}+1} \quad (2)$$

where n is the number of transferred electrons, F is the Faraday constant, c is concentration, D diffusion coefficient, r_0 is the radius of the electrode, and L is the value of the recess height. We rearranged the equation from [2] to calculate D:

$$D = \frac{I\left(\frac{4L}{\pi N r_p}+1\right)}{4nFcNr_p} \quad (3)$$

where N is the estimated number of pores and r_p represent the radius of single pore between NPs. Therefore the $r_0 = Nr_p$ because it is approximation of all pores in the surface area of electrode. The equation, which can be derived from Fig. 2D, for obtaining r_p is:

$$r_p = Rcos30° - R \quad (4)$$

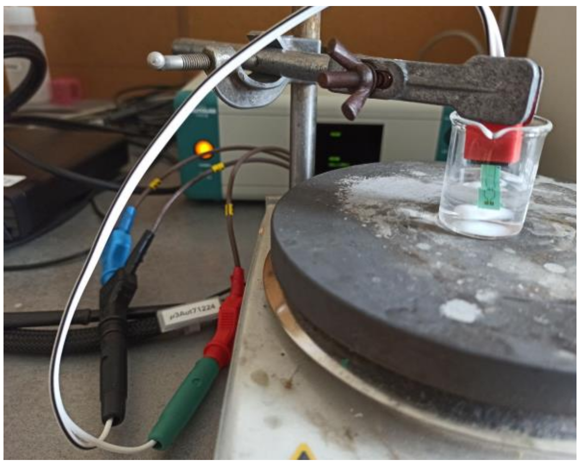


Fig. 1. Measurement aparature with biamperometric setup connected to µAUTOLAB III/ FRA2 (Metrohm Autolab, Netherlands) analyzer. Electrodes have circular shape, and are located in visible golden squares.

The possibility to approximate r_p this way has already been proposed in [4]. The R is the radius of NP. The relation is derived from the assumptions of ideal monolayers formed with NPs arranged closely together. The number of pores is calculated from:

$$N = \frac{S}{(2R)^2} \quad (5)$$

where S is area of the electrode $S = 0.378 \ 10^{-6}$ m.

B. Potential step chronoamperometry approach

Second principle to calculate the diffusion coefficient is to use potential step chronoamperometry approach [7][8]. It is a technique used in electrochemistry to study the kinetics of electrode reactions, particularly for studying diffusion-controlled electrode processes. The potential step is applied to an electrode, and the current response of the system is measured as a function of time. The current response is analyzed using the Cottrell equation [6]:

$$I = \frac{FS\sqrt{D}}{\sqrt{\pi t}}[A]_{bulk} \quad (6)$$

where D is the diffusion coefficient, $[A]_{bulk}$ the bulk concentration of species A in the solution, F is the Faraday constant and S the electrode area, t is the time, we can express the fitting function:

$$I = \frac{K}{\sqrt{t}} + C \qquad K = \frac{FS\sqrt{D}}{\sqrt{\pi}}[A]_{bulk} \quad (7)$$

The diffusion coefficient was calculated from K:

$$D = \left(\frac{K\sqrt{\pi}}{FS[A]_{bulk}}\right)^2 \quad (8)$$

TABLE I. LIST OF VARIABLES USED IN THE PAPER.

Variable	Description
$[A]_{bulk}$	the bulk concentration of species ***A*** in the solution
c	concentration
D	diffusion coefficient
F	Faraday constant
I	current
L	value of the recess height
n	number of transferred electrons
R	radius of NP
r_0	radius of the electrode
r_p	radius of single pore
S	area of the electrode
t	time

III. METHODS AND MATERIALS

A. Chemicals

Carboxylated polystyrene nanoparticles (PS NPs) of diameter ≈ 40 and 80 nm were purchased from Magsphere Inc. (U.S.A.). Poly-L-lysine hydrobromide (p-lys), N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC), N-Hydroxysuccinimide (NHS), glucose oxidase (GOD) from Aspergillus Niger, bovine serum albumin (BSA), Hexaammineruthenium(III) chloride (98%) (Ru^{3+}), potassium ferrocyanide and potassium ferricyanide (Fe^{2+}/Fe^{3+}), glutaraldehyde solution (GO) (25%), 2-(N-morpholino)ethanesulfonic acid (MES) were purchased from Sigma-Aldrich (Germany). D-(+)-Glucose monohydrate, KOH, H_2O_2 (30%), isopropyl alcohol was purchased from Penta (Czech Republic).

B. Aparature and electrodes

Chronoamperometry was recorded using µAUTOLAB III / FRA2 (Metrohm Autolab, Netherlands) analyzer, with help of NOVA software. The commercial electrodes in a biamperometric setup (www.printed.cz, Czech Republic) of a pair of two identical gold disk electrodes were used in this work. The electrode has internal diameter 400 µm and there is no separate reference or counter electrode.

C. Preparation of samples

The commercial electrodes in a biamperometric setup (www.printed.cz, Czech Republic) of a pair of two identical gold disk electrodes were used in this work. The electrode has internal diameter 400 µm and there is no separate reference or counter electrode. The aim was to immobilize polystyrene nanoparticles (PNPs) with negative charge, of diameters approximately 40 nm and 80 nm, onto a thin layer of positively charged poly-L-lysine (p-lys) on a gold electrode. To prepare the electrode, it was first activated by polishing with microcloth and isopropyl alcohol to remove any residual photoresist and passivation layers from the surface. It was then treated with a solution of 0.5 M KOH and 20% H_2O_2 for 10 min. Next, amino groups were introduced onto the gold surface by functionalizing the gold electrode with 50 µg/mL p-lysine in phosphate-buffered saline (PBS; 10 mM, pH 7.4) through adsorption for at least 30 minutes. Following washing and drying steps, and then the solution of 1% w/v PS NPs, dispersed by ultrasound for 20 min, with temperature lower than 35°C, was applied onto the electrode modified with positive charge and left for 40 minutes. Next objective was to immobilize the enzyme. Two different immobilization procedures of the enzyme GOD were compared. In both cases, the enzyme was applied only on one of the two electrodes. The first method involves cross-linking using a bi-functional agent, while the second method involves introducing biomolecules onto the surface of PNPs through covalent binding.

Crosslinking method (Fig. 2C): The mixture of 16 µl GOD (8 mg/ml) and 25 µl BSA (16 mg/ml) in the optimal concentrations to be cross-linked with 2.75 µl 2% glutaraldehyde (GO) solution was prepared. The microliter drop was applied to the PNPs modified electrodes and let dry. The reference sample was created as 33 µl BSA (13 mg/ml) with 2.75 µl 2% glutaraldehyde solution, without GOD. The second reference was the cross-linked GOD enzyme with BSA, but without NPs.

Covalent binding (Fig. 2A,B): Covalent binding of GOD to NPs was initiated by the activation of carboxy-groups of PNPs with a mixture of 5 mM EDC and 5 mM NHS dissolved in MES buffer (pH 6.2) for 15 min. After rinsing the electrode in MES buffer, the GOD (8 mg/ml) solution was applied to the electrode and incubated for 24 hours. Then the electrodes were rinsed in PBS buffer and used for the measurement. The reference sample was created by replacing enzyme GOD with protein BSA (8 mg/ml) and incubated for 24 hours. The second reference was the covalent bounded GOD enzyme to glutaraldehyde 2% (2.75 µl, 2%, 40 min), but without NPs.

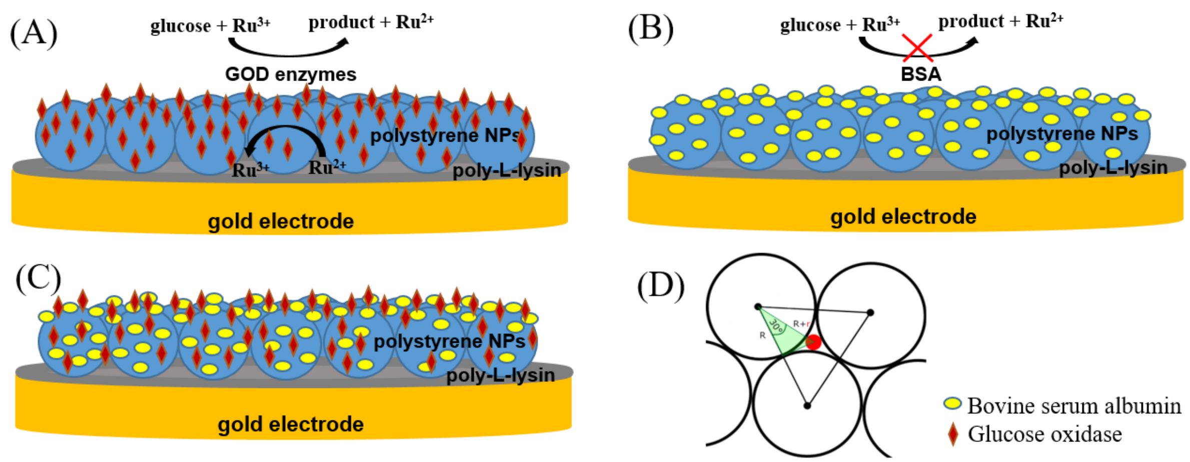
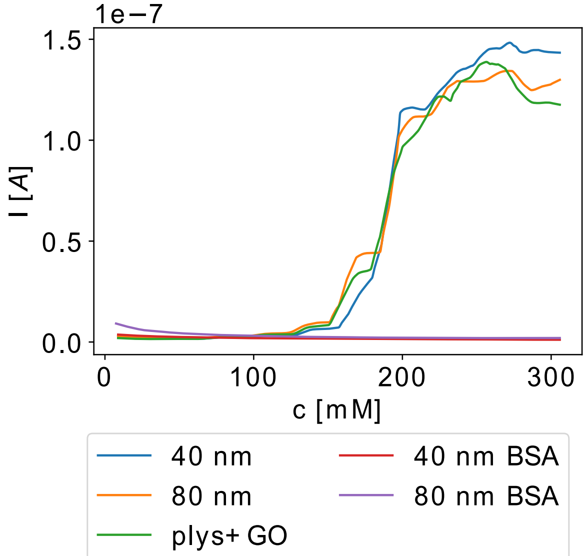


Fig. 2. Scheme of immobilization of GOD and BSA covalently (A and B) and by cross-linking method (C). The pore formation model and suggestion of pore size calculation (D).

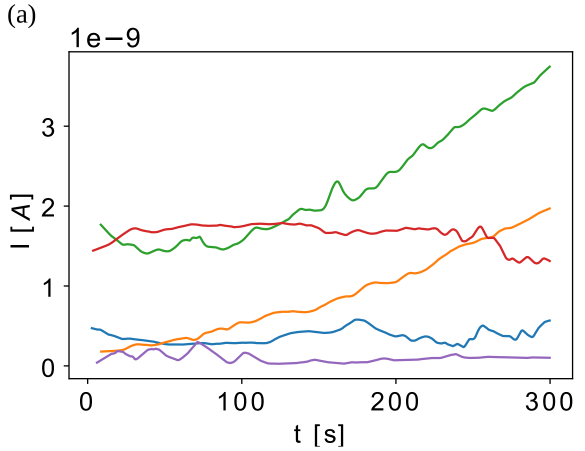
IV. Results and Discussions

To determine the diffusion coefficients, it was first necessary to perform amperometry. To get as much information as possible about the sensor, we performed calibration measurements, with gradually increasing glucose concentration (Fig. 3a). It also check correct function of this biosensor. Then, the most significant jumps of amperes were selected for further analysis of the diffusion coefficient. However, it turned out that although cross-linking of the enzyme was successful for these conditions, the same could not be said for covalent binding (Fig. 3b). The covalent bounding had not distinguishable distribution and characteristic only constantly increased. Therefore, only cross-linking samples were usable for this method and following text will deal more with them only.

At first, diffusion coefficient was calculated from equation (3). Current I was used from chronoamperometry as the difference between two currents for different concentration of glucose. As recess height L was used diameter of NPs. Diametres of pores was counted as: 6.2 nm and 12.4 nm. Number of pores (N) were approximated from equation (5) as 2×10^8 for 40 nm NPs and 5×10^7 for 80 nm NPs. Values for r_p counted from equation (4) was estimated as $r_{p(40)} = 3.1$ nm and $r_{p(80)} = 6.2$ nm. Diffuse coefficient for NPs of cross-linked method GOD bonding was calculated from averaged currents which are in Table II and results are listed in Table III. Numbers in brackets in Table III determine which equation was used to calculate them. Diffuse coefficient for the reference, electrode without NPs with crosslinked enzyme GOD on p-lys, was taken for counting D from equation (1), because there are no polystyrene nanopores. The sample with p-lys were chosen, because then we can observe just change created by the NPs. Assumption was that D would be lower for samples containing NPs because polystyrene particles shield the electrode, which for counting from equation (3) turn out to be true. According to literature the ruthenium(II) haves diffusion coefficient of about $D_{RuII} = 7.9\ 10^{-10}\ \frac{m^2}{s}$ [9] same as ruthenium(III) $D_{RuIII} = 7.9\ 10^{-10}\ \frac{m^2}{s}$ [10] and glucose in water $D_{Gl} = 6\ 10^{-10}\ \frac{m^2}{s}$ [11].



(a)



(b)

Fig. 3. Calibration chronoamperometry of the cross-linked (a) and covalent bounded (b) enzyme GOD samples

TABLE II. ABSOLUTE VALUES OF THE DIFFERENCE IN CURRENT [NA] CAUSED BY THE ADDITION OF GLUCOSE. THE VALUES ARE OBTAINED BY AVERAGING THE DIFFERENCE BETWEEN ADDICTION OF GLUCOSE FROM INDIVIDUAL MEASUREMENTS.

Glucose	**Step in current [nA] for glucose concentrations**		
	40 nm	***80 nm***	***Plys+GO***
0.75 -- 1 mM	3.5 ± 0.2	7.1 ± 0.6	5.9 ± 0.2
1.5 -- 3 mM	85 ± 13	67 ± 5	62 ± 5

For our experiment the meassurable diffusion is affected by traveling of glucose in liquid and by Ru(II) to the electrode not only from Ru(III), but as we see from literature, they diffuse coefficients are very similar, so the difference is neglected. Our values D without NPs are 100x higher, which may be caused by

TABLE III. DIFFUSION COEFFICIENTS [$\frac{m^2}{s}$] OF PLYS+GO AND PLYS+NPS COVERED ELECTRODES WITH GOD IMMOBILIZED BY CROSS-LINKING FROM CALIBRTION CHRONOAMPEROMETRIC MEASUREMENTS. NUMBERS IN BRACHIT SHOWS NUMBER OF EQUATION USED FOR CAUTING.

Glucose	**Diffusion coefficients from steady state current**				
	Dbare	***D_{40} NPs (3)***	***D_{80} NPs (3)***	***D_{40} NPs (2)***	***D_{80} NPs (2)***
1 mM	$4.4\ 10^{-8}$	$1.46\ 10^{-11}$	$1.48\ 10^{-10}$	$2.61\ 10^{-8}$	$5.31\ 10^{-8}$
10 mM	$1.54\ 10^{-7}$	$1.18\ 10^{-10}$	$4.67\ 10^{-10}$	$2.11\ 10^{-7}$	$1.69\ 10^{-7}$

the applied voltage, therefore the movement of particles is not free, but it is influenced in direction to the electrode. Secondly, the comparison with the potential step chronoamperometry approach for distinguish diffusion coeficient taken place. For measurement (Fig. 4) were used three different types of electrodes. The bear with cross-linked GOD, but this time without poly-lysine. Then GOD immobilized by the cross-linking method for 40 nm and 80 nm NPs. Covalent bonding was skipped because amperometry does not seem like reasonably sensitive method for this type of biosensor. Therefore, only samples with cross-linked enzyme were tested. The diffusion coefficient was retrieved by fitting the measured curves with equation (7) and then calculated from equation (8). The measured currents after potential increment from 0 mV to 0.5 mV were shown in Fig. 4. The diffusion coefficients are listed in Table IV. We can see that the electrode without a layer has a slightly lower value, but for a higher concentration of glucose, it is much lower. The products of glucose molecules may hinder movement and prevent a faster reaction process. In addition, since there is no complicated nanoterrain on the surface, where a larger amount of enzyme could be captured due to the increased surface area, this leads to a smaller number of ions arriving and transforming on the electrode, and therefore a diffusion coefficient is smaller. Between 40 NPs and 80 NPs, we can observe a difference in that for lower glucose concentrations (1 mM), D_{80} has twice the value, but at higher glucose concentrations, it is the opposite. It is possible that at low concentrations, 80 NPs benefit from the increased surface area for immobilization, while glucose products do not block their path. However, for higher concentrations, 40 NPs will benefit from a larger mobilization surface and balance out the poorer pore permeability. What is most important is how is it possible that the first and second method give such different trends between NPs and the non-porous electrode? If we used equation (2) instead of (3) (the values given in Table III labeled (2)) D would result into closer result. That is, if we did not consider all individual pores separately but took it as one surface and only included the height, the result would be closer between the methods. This seems more plausible, especially considering amperometry, where for lower concentrations, the 40 nm sample has lower currents than the other samples, while for higher concentrations, it is the opposite. That is consistent with equation (2) and potential step method. The equation (3) does not have to be completely wrong, but may be missing some necessary approximation, or our surface is not adequately porous.

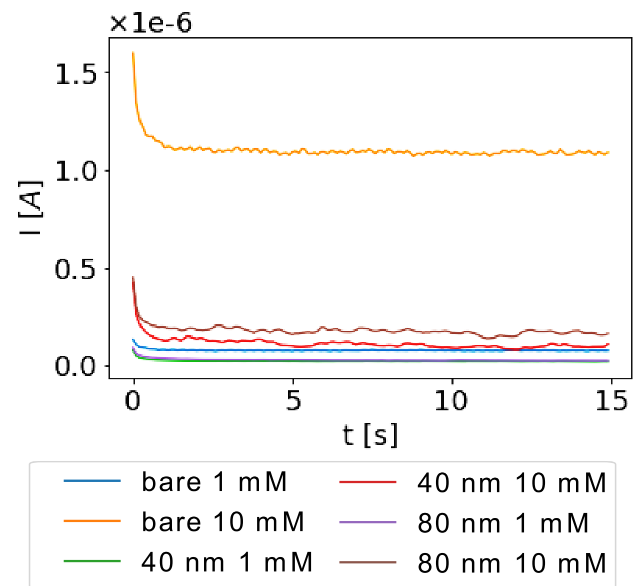


Fig. 4. Chronoamperometry after increment of potential from 0 to 0.5 mV. Measured separately for 1 mM and 10 mM concentration of glucose. Average of the curves were used to calculate the diffusion coefficient. Obtain only cross-linked samples.

TABLE IV. DIFFUSION COEFFICIENTS $[\frac{m^2}{s}]$ OF BARE ELECTRODES GOD AND NPS COVERED ELECTRODES WITH GOD IMMOBILIZED BY CROSS-LINKING FROM CHRONOAMPEROMETRIC MEASUREMENTS WITH POTENTIAL STEP FITTED BY COTTRELL EQUATION. THE UNCERTAINTY CONNECTED TO FITTING CURRENT IS AROUND 5 % OF VALUE.

Glucose	**Diffusion coefficients from Cottrell equation**		
	D_{bare}	$D_{40}NPs$	$D_{80}\ NPs$
1 mM	$1.22\ 10^{-7}$	$2.38\ 10^{-7}$	$4.34\ 10^{-7}$
10 mM	$1.59\ 10^{-10}$	$8.05\ 10^{-8}$	$4.48\ 10^{-8}$

V. CONCLUSION

We tested two different methods for the application of the GOD enzyme for use as an amperometric biosensor with a nanoporous matrix of polystyrene particles. It was found that while there is enough enzyme left on the electrode with cross-linking and we are able to measure the response for individual increases in glucose, for covalent bonding we get a continuous increase in current, but we are not able to measure individual additions, and this increase is only in the nA range. Therefore, further investigation of the diffusion coefficient focused on cross-linking. Using two different methods: steady-state current and potential step chronoamperometry, diffusion coefficients were calculated. In comparison with an empty electrode with cross-linking GOD, we can say that NPs in conjunction with p-lys affect diffusion to the electrode, but in potential step chronoamperometry, it was shown to improve rather than worsen it. This counterintuitive result is probably due to the fact that we used amperometry as a basis on which we built the measurement of incoming Ru ions, but by increasing the surface area for enzyme immobilization, more RuIII was necessarily produced, resulting in larger currents being measured. We also compared approximations from the modified expression for steady-state current. When not accounting for individual pores but considering only the height of the matrix, we obtained a result closer to the potential step method.

ACKNOWLEDGMENT

Created with financial support from project IGA – ID 8162 - Moderní mikro- a nanoelektronika pro budoucnostand project IGA – ID 8203 - Nanášení dielektrické vrstvy na wolframové elektrody a analýza jejich vlastností.

REFERENCES

[1] T. K. Perkins and O. C. Johnston, "A Review of Diffusion and Dispersion in Porous Media," *Soc. Pet. Eng. J.*, vol. 3, no. 01, pp. 70–84, Mar. 1963, doi: 10.2118/480-PA.

[2] J. V. M. and and P. R. Unwin*, "Determination of the Diffusion Coefficient of Hydrogen in Aqueous Solution Using Single and Double Potential Step Chronoamperometry at a Disk Ultramicroelectrode," *Anal. Chem.*, vol. 69, no. 12, pp. 2063–2069, 1997, doi: 10.1021/AC961211I.

[3] X. Yang *et al.*, "Mesoporous Materials–Based Electrochemical Biosensors from Enzymatic to Nonenzymatic," *Small*, vol. 17, no. 9, p. 1904022, Mar. 2021, doi: 10.1002/smll.201904022.

[4] J. Sopoušek, J. Humlíček, A. Hlaváček, V. Horáčková, P. Skládal, and K. Lacina, "Thick nanoporous matrices of polystyrene nanoparticles and their potential for electrochemical biosensing," *Electrochim. Acta*, vol. 368, p. 137607, Feb. 2021, doi: 10.1016/j.electacta.2020.137607.

[5] J. Sopoušek, J. Věžník, P. Skládal, and K. Lacina, "Blocking the Nanopores in a Layer of Nonconductive Nanoparticles: Dominant Effects Therein and Challenges for Electrochemical Impedimetric Biosensing," *ACS Appl. Mater. Interfaces*, vol. 12, no. 12, pp. 14620–14628, Mar. 2020, doi: 10.1021/acsami.0c02650.

[6] S. A. Nowinski and D. M. Anjo, "Determination of the Diffusion Coefficients of Catecholamines by Potential Step Chronoamperometry," *J. Chem. Eng. Data*, vol. 34, no. 3, pp. 265–268, Jul. 1989, doi: 10.1021/JE00057A002.

[7] N. Godino, X. Borrisé, F. X. Muñ, F. Javier Del Campo, and R. G. Compton, "Mass Transport to Nanoelectrode Arrays and Limitations of the Diffusion Domain Approach: Theory and Experiment", doi: 10.1021/jp9031354.

[8] O. V. Klymenko, R. G. Evans, C. Hardacre, I. B. Svir, and R. G. Compton, "Double potential step chronoamperometry at microdisk electrodes: simulating the case of unequal diffusion coefficients," *J. Electroanal. Chem.*, vol. 571, no. 2, pp. 211–221, Oct. 2004, doi: 10.1016/J.JELECHEM.2004.05.012.

[9] J. Moldenhauer, M. Meier, and D. W. Paul, "Rapid and Direct Determination of Diffusion Coefficients Using Microelectrode Arrays," *J. Electrochem. Soc.*, vol. 163, no. 8, pp. H672–H678, May 2016, doi: 10.1149/2.0561608jes.

[10] Y. Wang, J. G. Limon-Petersen, and R. G. Compton, "Measurement of the diffusion coefficients of [Ru(NH3) 6]3+ and [Ru(NH3)6]2+ in aqueous solution using microelectrode double potential step chronoamperometry," *J. Electroanal. Chem.*, vol. 652, no. 1–2, pp. 13–17, Mar. 2011, doi: 10.1016/j.jelechem.2010.12.011.

[11] P. Mohankumar, J. Ajayan, T. Mohanraj, and R. Yasodharan, "Recent developments in biosensors for healthcare and biomedical applications: A review," *Meas. J. Int. Meas. Confed.*, vol. 167, p. 108293, Jan. 2021, doi: 10.1016/j.measurement.2020.108293.

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

Influence of High Concentration of Silica Nanoparticles on the Dielectric Spectra

Ammar Al Soud
Central European Institute of Technology
Brno University of Technology
Purkynova 656/123,61200, Czech Rpublic
Ammar.al.soud@vutbr.cz

Samer Daradkeh
Central European Institute of Technology
Brno University of Technology
Purkynova 656/123, 61200 Brno, Czech Rpublic
samer.daradkeh@ceitec.vutbr.cz

Alexandr Knápek
Institute of Scientific Instruments of the Czech Academy of Sciences
Královopolská 147,612 64, 62100 Brno, Czech Rpublic
knapek@isibrno.cz

Karel Liedermann
Department of Physics
Brno University of Technology
Technická 24848/8,61600 Brno,Czech Republic
liederm@vut.cz

Vladimír Holcman
Department of Physics
Brno University of Technology,
Technická 2848/8, 61600 Brno, Czech Republic
holcman@vut.cz

Dinara Sobola
Central European Institute of Technology
Brno University of Technology
Purkynova 656/123, 61200 Brno, Czech Republic
sobola@vut.cz

***Abstract*— In the presented work, we report the dielectric behavior of epoxy–silicon oxide composites in the temperature range 240– 300 K, over the frequency range 10^{-2} Hz – 10^{7} Hz. The measuring apparatus was based on the Novocontrol Alpha Analyser and the measured data were analyzed and interpreted using the Havriliak – Negami equation. The master curves of the real part of permittivity and the dielectric loss number were obtained by time–temperature superposition principle, and the results showed that the nano-composite had a much higher loss factor. Through the analysis of the origin of the dielectric response in epoxy/silica composite, the reason for the different dielectric relaxation behaviors of the nano-composite, and the pure epoxy was discussed.**

***Keywords—nanofillers;* dielectric relaxation Spectroscopy, *epoxy resin*, nanocomposite.**

I. INTRODUCTION

Epoxy is one of the greatest options among the various types of polymers in particular applications where robust, fracture-resistant materials with high mechanical endurance are required [1].Moreover, epoxy resins have only recently been employed to improve the emission current properties of cold emission cathodes by creating a thin epoxy layer on their surface. Furthermore, epoxy resins find widespread applications in electrical engineering, particularly in high voltage systems. [2]–[5]. The relatively extensive usage of epoxy resin is attributed to its exceptional electrical, thermal, and mechanical properties, as well as its ease of formation. [4]. Epoxy has excellent dielectric properties in addition to being known for having the highest degree of mechanical endurance [1].

Enhancements in the dielectric and mechanical properties of polymers are being made possible by developments in the field of nanotechnology. Lately, composites based on different organic and inorganic fillers such as carbon [6], alumina oxide [7], titania oxide [8] are described with better dielectric characteristics as compared to pure epoxy. In a different study[1], [6], composites had a lower loss factor as a function of frequency than pure epoxy. Some studies have shown that nano- and micro-particles have a significant influence on the aging behavior of polymers [1], [8].

According to recent investigations, It has also been observed that the dielectric permittivity and the loss factor values of epoxy/nanofillers at lower concentrations (less than 0.5 wt%) are often lower than pure epoxy over a wide range of frequencies [9], [10]. Based on this observation, this work aims to study the effects of a high concentration of silicon dioxide nanoparticles (SNPs) (5 wt%, and 10 wt%) on the dielectric properties of epoxy/SNPs at different temperatures by comparing those results with the data obtained from the pure epoxy resin. In this context it was used Dielectric relaxation spectroscopy (DRS) for examining the electrical properties pure epoxy and epoxy nanocomposites [3]. The dielectric constant, loss number, conductivity, and dielectric relaxation behaviors have been reported based on the Havriliak-Negami (HN) equation.

II. Materials and Experimental Methods

A. Samples Preparation

The Epoxylite® E478 (E-478), purchased by Elantas (Wessel, Germany), was used to prepare the DRS samples. High-purity SNPs (99.5%) with a particle size range of (3 to 15) nm were obtained from Sigma-Aldrich, USA. The epoxy resin was mixed with the SNPs according to three steps. First, 5 g of epoxy resin was placed in a container, it was mixed with 2 ml of ethanol for 10 minutes in order to reduce the viscosity of the epoxy. In the second step, the epoxy was mixed with SNPs mechanically for 10 minutes. The following stage involved adding SNPs to the dilute epoxy and stirring it mechanically for 10 minutes. Also, the mixture container was placed in an ultrasonic bath for two hours to ensure arbitrary distribution of silica nanoparticles. The mixture was poured into silicone molds that are resistant to high temperatures. The molds had a depth of 3 mm and a radius of 5 cm. The samples were baked in two stages. First, the mixture was baked at 80 ℃ for 8 hours in order to remove air bubbles within the mixture and to evaporate the ethanol. In the second stage, the mixture was annealed at 180 ℃ for one hour to harden the mixture,[2], [11]. After the hardening process, the samples were polished using a semi-automated grinding/polishing machine (TEGRAMIN 30, Struers, Denmark) to have approximately 250 µm thin layer of composite epoxy layer. The samples thickness was measured with a micrometer. After the polishing process, the samples were cleaned ultrasonically in a bath of distilled water for 10 minutes. In summary, we have three samples, pure epoxy, epoxy with 5wt% of SNPs, and with 10wt% of SNPs.

III. results and discussion

A. Characterization

A.1. Fourier transform infrared spectra

Transmittance spectra of samples were acquired by vacuum Fourier transform infrared (FTIR) Vertex 70v, manufactured by (Bruker, USA). Figure 1 shows the FTIR spectra of pure epoxy and epoxy/ SNPs spectra, which display the presence of SiO_2, Si–OH, and O–H bending peak at 3433 cm^{-1} [12]. Also, C=O and C=C bonds 1640 cm^{-1}. And, the COOH is located at the pure epoxy spectra at 3630 cm^{-1} [13]. The characteristic C-OH stretching is responsible for the conspicuous absorption peak at 3590 cm^{-1}. The hydroxyl bond with the secondary carbon atom in the epoxy resin molecular backbone is confirmed by the absorption peak at 1590 cm^{-1}. The bands at 2130 cm^{-1} and 1890 cm^{-1} are caused by the vibrational stretching of carbon dioxide Co_2 and the methyl group. The band at 1470 cm^{-1} is caused by the vibrational stretching of bending of primary amine N–H. In the epoxy/SNPs spectra, it was observed that all peaks disappeared except for the peaks of SiO_2 at 3433 cm^{-1} and CH_2 at 1890 cm^{-1}[7], [12], [13].

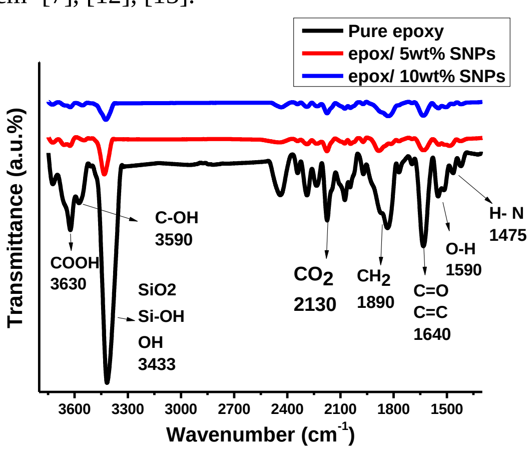


Fig. 1 FTIR spectra of pure epoxy and epoxy/SNPs.

A.2. X-ray photoelectron spectroscopy (XPS) spectra

The surface of epoxy/SNPs was chemically characterized using XPS (Kratos, UK) analysis to determine its elemental composition and chemical states. All sample surfaces contain silicon, carbon, nitrogen, and oxygen. CasaXPS software was used to precisely determine the composition of the functional groups of the Si and C atoms. The percentage of both carbon, oxygen, and nitrogen on the sarfece of pure epoxy sample was 78%, 17%, and 2%, respectively. While the percentage of silicon increased from 2% in the pure sample to 4% in the epoxy with 5wt% of SNPs sample, and 7% in the wt% SNPs sample. Fig. 2 shows the XPS spectra of pure epoxy and epoxy/5wt% SNPs. The main distinction between epoxy and pure epoxy/SNPs was seen in the corresponding high-resolution spectra of silicone (Si 2p), as illustrated in Fig.2. Pure epoxy's Si 2p spectrum shows two distinct peaks with FWHM values of 1.47 eV at 98.5 and 1.27 eV at 99.4 eV. These peaks are related to CH_2-SiO_3 and SiOH, respectively. While Si 2p spectra for epoxy/SNPs exhibit a high peak of 1.53 eV FWHM at 98.16 eV[12]. These differences in the Si 2p spectra of the pure epoxy and the epoxy/SNPs are evident from the increase of SiO_2 concentration within the epoxy. Moreover, the O 1s spectra show a peak of at 528.15 eV with 1.74 eV FWHM, further supporting the development of SiO_2 phases. The O 1s peak linked to O-Si bondsin SiO_2 emerges at 528.15 eV. Upon deconvolution of the O 1s spectra, the second peak with an FWHM value of 1.51 eV at 529 eV is attributed to the O-C and O-H bonds present in the epoxy backbone. Likewise, the C 1s spectrum was divided into three distinct parts. The C-C or C-H group is associated with the first C 1s peak, which is located at 284 eV. The second peak, which is associated with C-atoms in the C-O-C and C-O-H groups, is located at 285.51 eV [9]. The third peak in the C 1s spectrum, positioned at 288.22 eV, can be attributed to the presence of C=O doule bound, which may exist as oxidized polymer components or adsorbed CO2. The high-resolution N 1s spectrum is a valuable tool for XPS characterization of epoxy/SNPs.

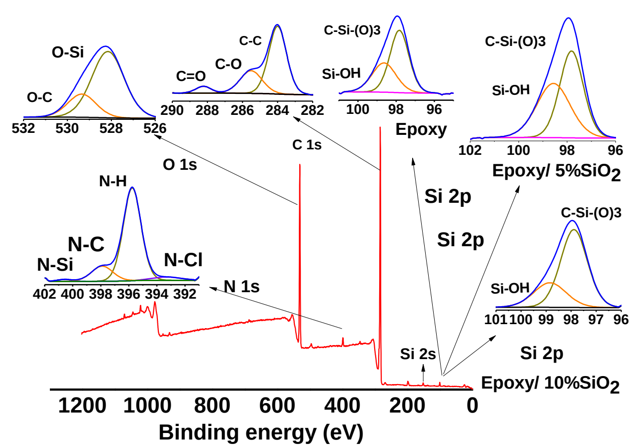


Fig. 2 The XPS spectra of pure epoxy and epoxy/SNPs.

The single N 1s peak at 395.47 eV is attributable to triamine. A singular N 1s peak is indicative of crosslinking between the epoxy and amine monomers [1], [9], [12].

B. DRS results

Dielectric measurements were carried out on the Novocontrol Alpha-A analyzer with Quatro cryosystem (Novocontrol Technologies GmbH & Co. KG, Germany). The analyzer is capable to test a wide frequency range of 10^{-2} Hz – 10^{7} Hz. The duration of a single sweep did not exceed 15 ms and the measurement across the full frequency range with 63 measuring points takes about 35 minutes. The samples were investigated at three different temperatures 240 K, 270 K, and 300 K. All samples and the electrode diameter of 20 mm with a voltage of 3 V_{RMS} were used. The Havriliak-Negami (HN) theory was used to fit the recoded experimental data [14]. The HN equation is composed of the Cole-Cole and Cole-Davidson functions [15], which, separately, describe the symmetrical and uneven broadenings of the dielectric function. The HN equation for the complex permittivity is the most commonly used function for describing the dielectric properties of materials and their relaxation behavior in the frequency domain. which can be written as follows [14], [16]:

$$\hat{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{[1+(j\omega\tau)^\alpha]^\beta} \quad (1)$$

Where $\hat{\varepsilon}$ the complex permittivity, $\Delta\varepsilon = \varepsilon_s - \varepsilon_\infty$, ε_∞ is the permittivity at infinite frequency also called optical permittivity, ε_s is the static permittivity at zero frequency, α describes the flatness/width of the maximum, β expresses the skewness (asymmetry) of the complex dielectric permittivity. α and β it and be in range of ($0 < \beta \leq 1$ and $0 < \beta \cdot \alpha \leq 1$), these parameters show the changes in the relaxation time distribution. ω is the angular frequency and τ relaxation time indicates the position of the relaxation maximum on the horizontal axis. Table 1 shows, the parameters $\Delta\varepsilon$, τ, α, and β· related to the HN equation.The complex permittivity can be decomposed into the real ε′(ω) and imaginary ε″(ω) parts as follows [14], [16]:

$$\varepsilon'(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon\, cos(\beta\emptyset)}{\left[1+2(\omega\tau)^\alpha cos\left(\frac{\pi\alpha}{2}\right)+(\omega\tau)^{2\alpha}\right]^{\beta/2}} \quad (2)$$

$$\varepsilon''(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon\, sin(\beta\emptyset)}{\left[1+2(\omega\tau)^\alpha cos\left(\frac{\pi\alpha}{2}\right)+(\omega\tau)^{2\alpha}\right]^{\beta/2}} \quad (3)$$

Fig.3 shows the changes in permittivity and loss values in response to frequency and temperature changes. It can be seen in Fig 3, a, b, and c (left) that the value of the permittivity of the epoxy/ 5 wt% of SNPs sample was lower than that of the pure sample, up to 300 K. While the value of the permittivity of the epoxy/ 10 wt% of SNPs sample increased significantly with the increase in temperature. Based on this result, it can be said that the permittivity of epoxy/SNPs is lower than that of pure epoxy, provided that the filler concentrations are less than 5 wt% at room temperature lower. However, epoxy/ 5 wt.% of SNPs showed a decrease in loss value more than other samples at different temperatures, as shown in Fig.3 a, b, and c (right).
The changes in permittivity with the presence of nanofillers and their loading with high concentration can be summarized into two reasons: the accumulation of space charges within the resin system induces interfacial polarization between the base resin and nanoparticles. Only, interfacial polarization occurs under lower frequency ranges up to 10^3 Hz [17]. As a result, the permittivity decreases as an effect of interfacial polarization, especially with lower concentration, so it can be neglected.
Another reason is that the polarization of the epoxynanocompositewill decrease when the mobility of the dipole groups decreases.However,when the frequency of the applied field decreases, the Epoxy/ SNPs composites likewise experience a rise in active dangling bonds on the surface of the nanoparticles, which results in a decrease in electric permittivity of the epoxy [9]. In more detail, the addition of a small quantity of nanoparticles to the epoxy leads to the formation of easily polarizable bonds under the influence of an electric field. However, as the amount of nanoparticles in epoxy increases, a numerous thin immobile nanolayers formed as a result of powerful interaction between the SNPs and polar groups in epoxy. The presence of these immobile nanolayers may ultimately limit the mobility of the side- or end-chain of epoxy [9], this explains the reason for the decreasing permittivity of the epoxy/ 5wt% SNPs sample. As the filler content rises, the high inherent permittivity of SNPs begins to have an impact on the relative permittivity of nanocomposites, this is known as Lichtenecker-Rother law [17]. This explains why the sample permittivity for epoxy/10 wt% SNPs is increased. Yet as the temperature increases to 300 Kelvin or higher, the movement of the dipole groups and the epoxy chain will accelerate, which increases the electrical permittivity.

Figure 4 illustrates the impact of SNPs on conductivity at various temperatures. As demonstrated in Fig. 4a, the conductivity reached its maximum values at 4.2×10^6 Hz, after which it began to decline. Similarly, Fig. 4b shows that the conductivity of epoxy/SNPs sample reached its maximum values at 6.2×10^6 Hz, after which it began to decline, In contrast, the conductivity rose for both other samples over the whole frequency range. As can be seen in Fig. 4c, the conductivity increased throughout the whole frequency range for all samples.

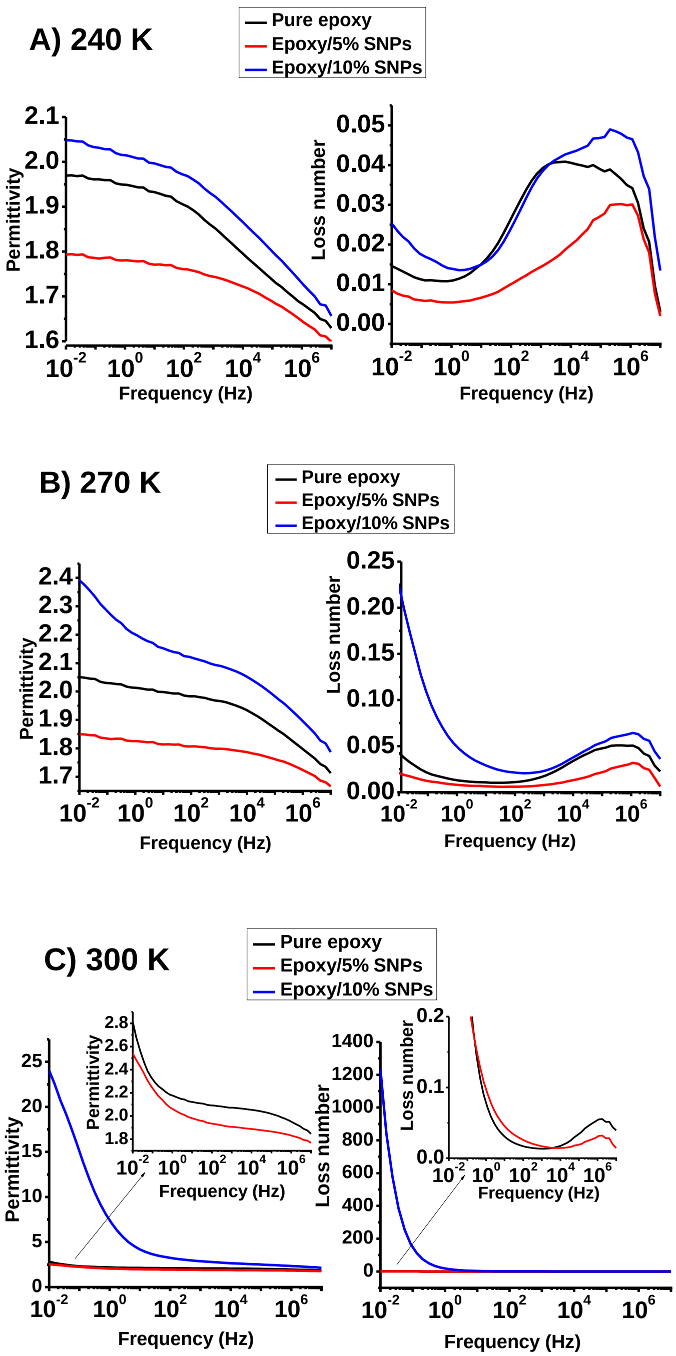


Fig. 3 Variations of permittivity(left) and loss number (right) with respect to frequency and temperature.

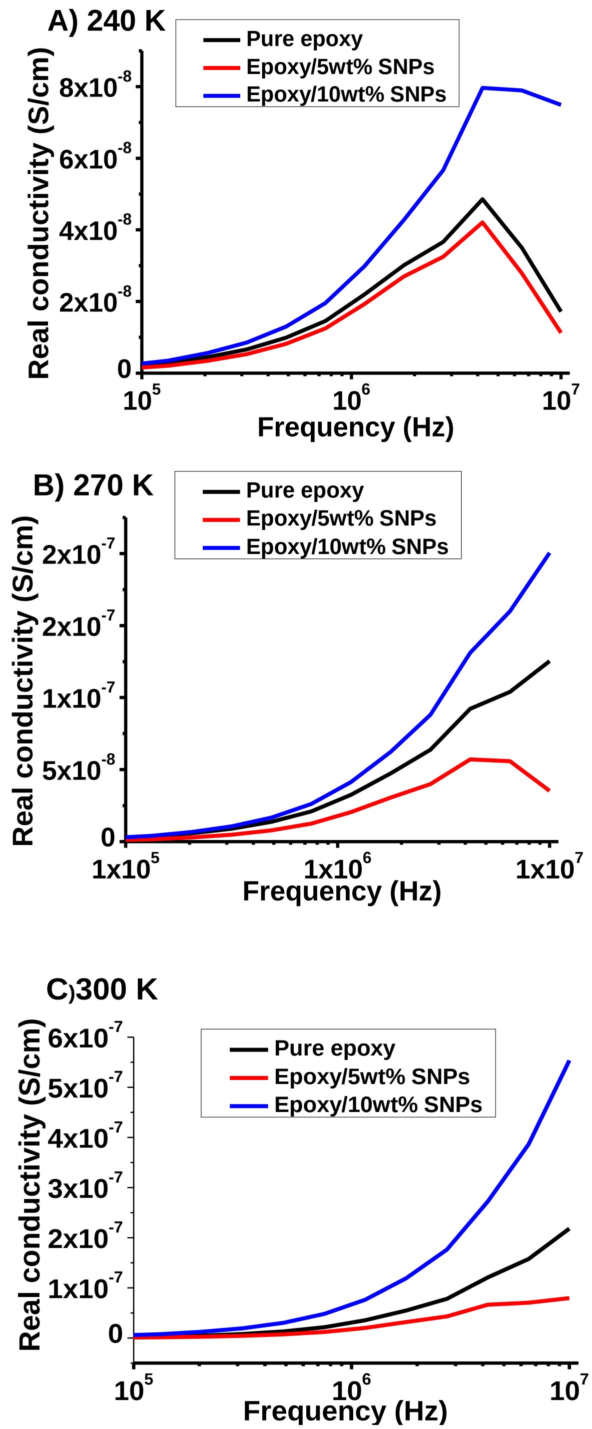


Fig. 4 Variations of conductivity with respect to frequency and temperature.

Table1:The conductivity and the HN equation parameters

Material	T (K)	$\Delta\varepsilon$	$\tau\,(s)$	α	β
Pure epoxy	240	0.353	1.1×10^{-5}	0.304	1
	270	0.34	5.9×10^{-7}	0.413	1
	300	0.84	1.05×10^{-8}	0.522	1
Epoxy/ 5%SiO_2	240	0.44	2.32×10^{-5}	0.725	1
	270	0.53	1.33×10^{-7}	0.811	1
	300	0.78	5.23×10^{-8}	0.872	1

IV. Conclusions

Dielectric spectroscopy has been used to examine the dielectric permittivity and loss number of pure epoxy and epoxy/SiO_2 nanoparticles loaded with high concentrations of SiO_2. Based on this investigation, the following conclusion has been drawn. At room temperature or lower, the epoxy/ 5 wt% SiO_2 sample has a lower permittivity, loss number, and conductivity than the pure epoxy sample. For epoxy/ 10 wt% SiO_2 sample has a higher permittivity, loss number, and conductivity than the pure epoxy sample. The reason for this is when SNPs are added to an epoxy, thin immobile nanolayers are created inside the resin that impedes the movement of the dipoles and the chain of the epoxy.

ACKNOWLEDGMENT

The Czech Nano Lab Research Infrastructure supported by MEYS CR (LM2018110) is acknowledged. We also acknowledge the Czech Academy of Sciences (RVO: 68081731) for providing the etching equipment and laboratory infrastructure.

REFERENCE

[1] A. Khattak *et al.*, “Effects of Compression and Silica Addition on the Dielectric Properties of Epoxy Composites,” *Arab J Sci Eng*, vol. 45, no. 8, pp. 6741–6750, Aug. 2020.

[2] A. Al Soud, R. N. Al Buqain, and M. S. Mousa, “Composite Metallic Nano Emitters Coated with a Layer of Insulator Covered by Au layer,” Jordan Journal of Physics, vol. 13, no. 3, pp. 253–262, 2020.

[3] M. Klampar, M. Spohner, P. Skarvada, D. Dallaeva, J. Kobrtck, and K. Liedermann, “2013 IEEE Electrical Insulation Conference, EIC 2013,” 2013, pp. 286–290.

[4] A. Al Soud, A. Knápek, and M. S. Mousa, “Analysis of the Various Effects of Coating W Tips with Dielectric Epoxylite 478 Resin or UPR-4 Resin Coatings under Similar Operational Conditions ,” Jordan Journal of Physics, vol. 13, no. 3, pp. 191–199, 2020.

[5] L. Chen and M. M. El-Gomati, “Field emission studies of tungsten-coated silicon-based "eld emitters,” 1999.

[6] D. Meisak, J. Macutkevic, A. Selskis, P. Kuzhir, and J. Banys, “Dielectric relaxation spectroscopy and synergy effects in epoxy/mwcnt/ni@c composites,” *Nanomaterials*, vol. 11, no. 2, pp. 1–18, Feb. 2021.

[7] B. Zhang, W. Gao, P. Chu, Z. Zhang, and G. Zhang, “Trap-modulated carrier transport tailors the dielectric properties of alumina/epoxy nanocomposites,” *Journal of Materials Science: Materials in Electronics*, vol. 29, no. 3, pp. 1964–1974, Feb. 2018.

[8] W. Jilani *et al.*, “DGEBA epoxy polymer composite panels (MB: DGEBA-EPCPs): structural, optical, and dielectric properties-leanings and appliances in optical electronics,” *Opt Quantum Electron*, vol. 55, no. 3, Mar. 2023.

[9] N. Phonthamachai, H. Chia, X. Li, F. Wang, W. W. Tjiu, and C. He, “Solvent-Free One-Pot Synthesis of high performance silica/epoxy nanocomposites,” *Polymer (Guildf)*, vol. 51, no. 23, pp. 5377–5384, Oct. 2010.

[10] Y. S. Lin, Y. H. Liao, and C. H. Hu, “Effects of N2 addition on enhanced scratch resistance of flexible polycarbonate substrates by low temperature plasma-polymerized organo-silicon oxynitride,” *J Non Cryst Solids*, vol. 355, no. 3, pp. 182–192, Feb. 2009.

[11] Q. Wang, G. Chen, and A. S. Alghamdi, “Proceedings of the 2010 IEEE International Conference on Solid Dielectrics, ICSD 2010,” 2010, pp. 4–9.

[12] A. Afzal, H. M. Siddiqi, S. Saeed, and Z. Ahmad, “Exploring resin viscosity effects in solventless processing of nano-SiO2/epoxy polymer hybrids,” *RSC Adv*, vol. 3, no. 12, pp. 3885–3892, Mar. 2013.

[13] A. Radoń, P. Włodarczyk, A. Drygała, and D. Łukowiec, “Electrical properties of epoxy nanocomposites containing Fe3O4 nanoparticles and Fe3O4 nanoparticles deposited on the surface of electrochemically exfoliated and oxidized graphite,” *Appl Surf Sci*, vol. 474, pp. 66–77, Apr. 2019.

[14] S. Havr1liak and S. Negami, “A Complex Plane Representation of Dielectric and Mechanical Relaxation Processes in Some Polymers.”

[15] D. W. Davidson and R. H. Cole, “Dielectric relaxation in glycerine [11],” *The Journal of Chemical Physics*, vol. 18, no. 10. p. 1417, 1950.

[16] I. Soroceanu *et al.*, “Broad-Band Dielectric Spectroscopy Reveals Peak Values of Conductivity and Permittivity Switching upon Spin Crossover,” *Journal of Physical Chemistry Letters*, vol. 10, no. 23, pp. 7391–7396, Dec. 2019.

[17] H. Shi, N. Gao, H. Jin, G. Zhang, and Z. Peng, “Investigation of the effects of nano-filler on dielectric properties of epoxy based composites,” 2009, pp. 804–807.

A Comparison Particle Filter for Searching a Radiation Source in Real and Simulated World

1st Milos Cihlar
Dept. of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
ORCID: 0000-0002-5478-9196

2nd Tomas Lazna
Central European Institute of Technology
Brno University of Technology
Brno, Czech Republic
ORCID: 0000-0001-5437-4014

3nd Ludek Zalud
Dept. of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
ORCID: 0000-0003-2993-7772

***Abstract*—In this paper, we are focusing on comparing solutions for localizing an unknown radiation source in both a Gazebo simulator and the real world. A proper simulation of the environment, sensors, and radiation source can significantly reduce the development time of robotic algorithms. We proposed a simple sampling importance resampling (SIR) particle filter. To verify its effectiveness and similarities, we first tested the algorithm's performance in the real world and then in the Gazebo simulator. In experiment, we used a 2-inch NaI(Tl) radiation detector and radiation source Cesium 137 with an activity of 330 Mbq. We compared the algorithm process using the evolution of information entropy, variance, and Kullback-Leibler divergence. The proposed metrics demonstrated the similarity between the simulator and the real world, providing valuable insights to improve and facilitate further development of radiation search and mapping algorithms.**

***Index Terms*—Particle filter, Particle filter comparison, Simulation of radioactivity, Kullback-Leibler divergence, Information entropy, Autonomous radiation search**

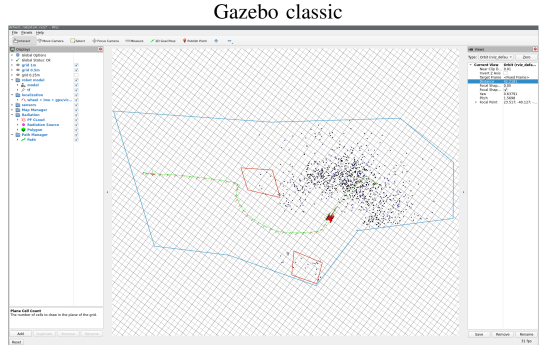
I. Introduction

Nuclear technology has a crucial role in numerous industries, medical, agricultural, and nuclear power plants. The use of radiation has brought many beneficial usages; however, it can also cause a significant risks to both human health and the environment. Environmental contamination can result not only from accidental radiation leaks but also from intentional misuse of nuclear materials by terrorist organizations, warfare, or damage to a nuclear power plant. Since 1995, the Incident and Trafficking Database (ITDB) of the International Atomic Energy Agency (IAEA) has documented over 3,500 incidents involving the loss or theft of radioactive materials [1].

The potential consequences of nuclear incidents emphasize the importance of effective, safety and fast radiation measurement and serching. Due to the effects of radiation on human health, it is valuable to develop algorithms for finding lost radiation sources with unmaned ground vehicle (UGV) and unmanned aerial vehicle (UAV).

Estimating the position of a radiation source is challenging due to the highly nonlinear model and the lack of directional sensitivity with only one detector. Developing a searching algorithm requires a radiation source for evaluation. Experiments involving radioactive substances are demanding in terms of legislation, time, and organization. In recent decades, most organizations involved in robotics research and development have used simulators [2]. Proper simulation can reduce the time required for developing new algorithms. Therefore, this paper proposes a method for searching for a radiation source using a particle filter and focuses on comparing the proposed algorithm in both real-world and Gazebo simulation scenarios.

(a) An environment with UGV and radiation source in Gazebo classic



(b) Visualization particle filter algorithm in RViz2

Fig. 1: UGV searches for a lost radiation source

A. Contribution

Robotics simulators contribute to making the development of autonomous robots easier, faster, and cheaper, making it advisable to use them. By comparing simulated and real experiments, it becomes possible to detect differences between the model and reality. This knowledge can then facilitate the transfer of simulated problems to real hardware.

B. Outline

The paper is organized as follows: Section II describes the particle filter and the radiation model used in proposed method. We also provide a detailed description our proposed method for searching radiation sources and implementation details. Section III presents an overview of performed experiments and defines the metrics used to compare particle clouds. The experiment results are presented in a subsection III-B. Finally, we present our conclusions in the last section IV.

II. Particle filter

The particle filter is a suboptimal variant of a nonlinear Bayesian filter. The Bayesian filter seeks to construct the posterior probability density function (pdf) of the state based on knowledge of the system (1) and measurement (2) model [3].

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \tag{1}$$

The system state $\mathbf{x}_k$ is given generally by nonlinear function $\mathbf{f}_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_x}$ where $\mathbf{v}_{k-1}$ is process noise.

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k) \tag{2}$$

Similarly, the model state $\mathbf{z}_k$ is also given by nonlinear function $\mathbf{h}_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_n} \to \mathbb{R}^{n_z}$ where $\mathbf{n}_k$ is measurement noise.

The pdf is obtained from prediction (3) and update (4) step [3] [4] [5].

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1} \tag{3}$$

The prior pdf $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ is obtained from the system model (1) and the pdf $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ from the previous $(k-1)$ update step (4).

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \tag{4}$$

In the update step (4), the measurement $\mathbf{z}_k$ is used to correct the prediction $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$. The normalization constant $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$, which depends on the system model's state equation (2), is detailed in [3].

The predict and update equations provide a conceptual optimal solution to the nonlinear Bayesian problem. However, in many cases, the solution cannot be determined analytically. Therefore, the problem is solved with more restricted optimal Kalman and gradient-based filters. In cases where these restrictions are not met, suboptimal recursive solutions like the particle filter are used. [3]

A. Description

The particle filter is a sequential Monte Carlo method (SMC) that recursively estimates the hidden true state vector. The key aspect of the particle filter is representing the posterior pdf (5) with a finite set of random samples $\{\mathbf{x}^i_{0:k}, w^i_k\}^{N_s}_{i=1}$ associated with weights w. Each sample represents a possible solution of the problem with a probability determined by its weight. The number of samples N_s varies depending on the application, but it can become very large. If the $N_s \to \infty$, the samples describe the true posterior density. [6] [7]

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) \approx \sum_{i=1}^{N_s} w^i_k \delta(\mathbf{x}_{0:k} - \mathbf{x}^i_{0:k}) \tag{5}$$

A sampling importance resampling (SIR) algorithm was chosen for estimating the position and activity of a lost radiation source. This algorithm requires knowledge of the system dynamics (1) and measurement model (2). The probability $p(\mathbf{z}_k|\mathbf{x}^i_k)$ depends on the weight (6) of the particle.

$$w^i_k \propto p(\mathbf{z}_k|\mathbf{x}^i_k) \tag{6}$$

To avoid the degeneracy problem [7], the low variance resampling algorithm was chosen. This algorithm eliminates particles with low weight and replaces them with newly generated particles. The details of the resampling algorithm are described in the II-C section. The weights of the particles must satisfy the following statement (7). [3]

$$\sum_{i=1}^{N_s} w^i_k = 1 \tag{7}$$

Therefore, before resampling, the weights must be normalized (8) to ensure that the total weight is equal to one.

$$\mathbf{w^n_k} = (\sum_{i=1}^{N_s} w^i_k)^{-1}\mathbf{w_k} \tag{8}$$

B. Radiation model

Unstable atoms in radioactive materials emit various particles, such as alpha particles, beta particles, X-rays, and gamma particles. Radioactive decay is a stochastic and strongly nonlinear process that follows a Poisson probability distribution. Radiation detectors count the number of particles that have passed through the sensor per unit time, typically measured in counts per second (CPS). The CPS value is proportional to the inverse square law, which describes how the intensity of radiation decreases relative to the distance from the source, as shown in the equation. [8]

$$\lambda_T \propto \frac{CPS_0}{(x_r - x_0)^2 + (y_r - y_0)^2 + h} + Bg \tag{9}$$

The λ_T represents the theoretical value of the detector at a given position (x_r, y_r) and height (h) above the terrain when the radiation source emits CPS particles per second at position (x_0, y_0). The attenuation factor of air is neglected, and the natural radiation background is denoted as Bg.

Radiation detectors cannot count all particles because it is saturated by the detector's dead time a measured in seconds. Therefore, the theoretical value λ_T does not correspond to the measured value from the detector. Measured value λ can be modeled according to the equation (10). [9]

$$\lambda = \lambda_T e^{-a\lambda_T} \tag{10}$$

C. Implementation

This section provides a detailed description of the particle filter used in this work. A particle filter is a recursive Bayesian algorithm consisting of the following steps: initialization, measurement, correction, and resampling. Since the static position of the radiation source is assumed, the prediction step is omitted.

The initialization step involves uniformly distributing N_s particles with weight w within the desired area.

$$\mathbf{P} = \{\mathbf{p}^i, w^i\}_i^{N_s} \tag{11}$$

The particle set $\mathbf{P}$ is restricted to a 2D surface since the radiation source lies on the ground. Each particle $\mathbf{p}$ consists of uniformly distributed $x \sim \mathcal{U}(x_{min}, x_{max})$ and $y \sim \mathcal{U}(y_{min}, y_{max})$ coordinates, as well as activity $A \sim \Gamma(\alpha = 2, \beta = 6000)$ with a gamma distribution. The radiation background $Bg \sim \mathcal{U}(Bg_{min}, Bg_{max})$ is generated uniformly with low and high limits.

$$\mathbf{p} = \begin{pmatrix} x \\ y \\ A \\ Bg \end{pmatrix} \tag{12}$$

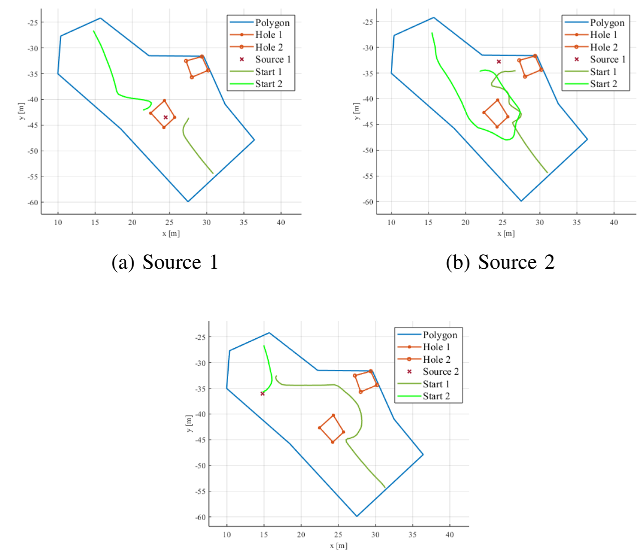
Measurements are taken at a frequency of $1Hz$ and include radiation values CPS and the position of the robot $\mathbf{x_r}$. In the correction step, the weights of each particle are computed according to its probability. The low variance resampling method [6] removes particles with low probability and replaces them with a new set of particles. The details of the particle filter are shown in Algorithm 1.

Algorithm 1 Particle Filter

```
Initialize : p = Random()
while var_T(P) ≤ const_1 do
    Measurement : z ← CPS
    for  i = 0 < N do
        λ_T = A / ((x_r − x)^2 + (y_r − y)^2 + h) + Bg
        λ = λ_T e^(−aλ_T)
        Weight: w(i) = N(z, λ, const_2 ∗ λ)
    end for
    Normalize : w_k = (Σ_{i=1}^{N_s} w_k^i)^{−1} w_k
    Low_variance_resampling_algorithm()
end while
```

III. Experiment

The particle filter was tested in real world and simulation experiment. In the experiment, a radiation source, Caesium-137 with 330 Mbq, was placed inside a polygonal area of over $500m^2$, which also had two restricted areas labeled as 'Hole1' and 'Hole2' as shown in the figure 2. Robot starts from two positions and search one radiation source placed in trhee different places in six experiments, see the table I. The robot was equipped by an RTK GNSS receiver and a 2-inch NaI(Tl) radiation detector with an estimated dead time of $a = 47 \times 10^{-5}s$. The initial start positions of the robot and the position of radiation sources are listed in the table I for each experiment. The trajectories were recorded, allowing the experiment to be repeated in the simulator.



(a) Source 1 (b) Source 2

(c) Source 3

Fig. 2: Experimental setup showing restricted area, radiation sources, and trajectories recorded from real-world experiment

TABLE I: Experiments overview

	Source [x [m], y [m]]	Start [x [m], y [m]]
Experiment 1	Source 1 = [24.45, -43.50]	Start 1 = [30.88, -54.47]
Experiment 2	Source 1 = [24.45, -43.50]	Start 2 = [14.74, -26.61]
Experiment 3	Source 2 = [24.46, -32.77]	Start 1 = [31.05, -54.43]
Experiment 4	Source 2 = [24.46, -32.77]	Start 2 = [15.44, -27.10]
Experiment 5	Source 3 = [14.76, -36.06]	Start 1 = [31.30, -54.38]
Experiment 6	Source 3 = [14.76, -36.06]	Start 2 = [14.95, -26.62]

The same experiments were conducted in Gazebo simulation Fig. 1. Each experiment was repeated five times, and the robot followed the same trajectories as in the real-world experiments. For simulation, the radiation sensor and detectors used the same Gazebo plugins as described in [9].

A. Evaluation methods

The generation of particles for filter subjects to a random process. Thus, comparing the evolution of two runs of the same algorithm cannot be accomplished by comparing their estimations with the true value. In this paper, the focus is on evaluating the algorithm's evolution, not just the result. The first method evaluates the total variance of the particles at each updating step.

$$var_T(\mathbf{P}) = \sqrt{cov(\mathbf{x}, \mathbf{x})^2 + cov(\mathbf{y}, \mathbf{y})^2} \tag{13}$$

The next one is the information (Shannon) entropy of particles computed at each iteration. It represents the average amount of information. If the entropy decreases to zero, the particles represent an estimation with no uncertainty. The logarithm (14) has a base 2; therefore, the information entropy is computed in bits in this paper. The information entropy for random events x is computed as follows (14);

$$H(x) = -\sum p(x)log_2(p(x)) \tag{14}$$

For multivariate distributed position of particles $\mathbf{P_p}$, the entropy $H(\mathbf{P_p})$ is computed with the following equation (15). The assumption is that the particles $\mathbf{P_p}$ are Gaussian distributed.

$$H(\mathbf{P_p}) = 1 + log(2\pi) + \frac{1}{2}log(|\mathbf{\Sigma}|) \tag{15}$$

$\mathbf{\Sigma}$ is the position covariance matrix of particles.

Kullback-Leibler divergence was used to compare two multivariate probability density functions $P : p_s \sim \mathcal{N}(\boldsymbol{\mu}_1, \mathbf{\Sigma_1})$ and $Q : p_r \sim \mathcal{N}(\boldsymbol{\mu}_2, \mathbf{\Sigma_2})$ represented with simulated resp. real particle clouds. [11] [12]

$$\begin{aligned} D_{KL}(P||Q) = \frac{1}{2}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \mathbf{\Sigma}_2^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) \\ + \frac{1}{2}tr(\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1) - \frac{1}{2}ln\left(\frac{|\mathbf{\Sigma}_1|}{|\mathbf{\Sigma}_2|}\right) - 1 \end{aligned} \tag{16}$$

The Kullback-Leibler divergence compares two probability density distributions and measures how P differs from the reference distribution Q.

B. Results

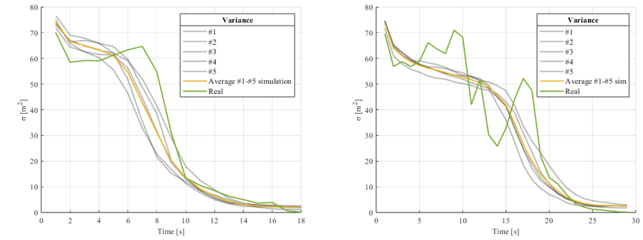
This section presents the results of our experiments, which we evaluated using three methods as described in the previous section III-A. We compared the results of the particle cloud obtained from the real-world experiment with the cloud obtained from the simulated experiments. Each simulated experiment was run five times with different source and position settings, see the table I. All data captured from the experiments were processed and divided into figures based on the position of the radiation source, and sub-figures were created based on different starting poses. For example, the left figure 3a displays data from the experiment with source 1 and start position 1, while the right figure 3b displays data from the experiment with source 1 and start position 2.

The following three figures 3 - 5 depict the variances of particles. The black transparent data represents the variances of simulated particles, denoted with a hash and number, while the orange line shows their average value. The green line represents the variances observed in the real-world experiment.

The particle filter variance is similar to the real world in Experiments 1, 3, and 6. In Experiments 2 and 5, true variance oscillates around the average variance of the simulated particles. But, the variance trend corresponds to the simulated particles variance. This difference may be caused by the different noise characteristics of the real and simulated detectors. Only the Experiment 4 is real-world experiment that does not match the simulation at all. This discrepancy is due to the simulated particle filter converges earlier than the real one. Additionally, the particles diffuse outside the polygon, leading to a higher variance than at the start. At 35 seconds into Experiment 4, Figure 6 shows the estimated probability density functions of both the real and simulated particle filters. The fast convergence of the simulated particle filter is clearly visible, in contrast to the real experiment where the particles have diffused over a much larger area.

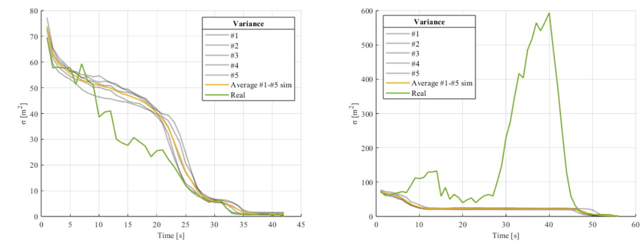
The information entropy evaluation is the most crucial evaluation method for particle filters in this paper. Entropy quantifies the amount of information contained in a random event, and the aim of the particle filter is to reduce the information entropy to zero in an ideal case. A probability distribution function with low entropy is highly skewed, and therefore the estimation falls into a smaller area with a high probability.

The similarities are visible in all experiments except for the fourth. The information entropy rate descent is very similar with the simulated experiment. The biggest difference



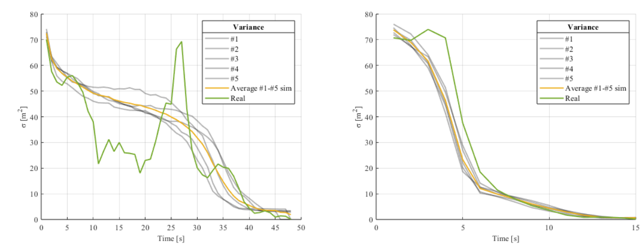
(a) Start 1 (Exp. 1) (b) Start 2 (Exp. 2)

Fig. 3: Source 1: Variance



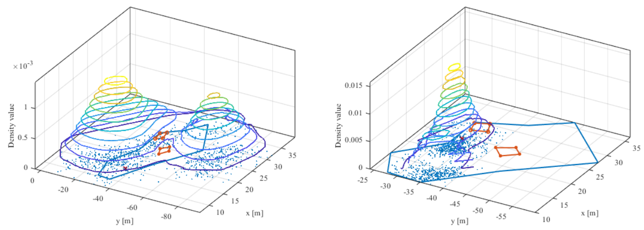
(a) Start 1 (Exp. 3) (b) Start 2 (Exp. 4)

Fig. 4: Source 2: Variance



(a) Start 1 (Exp. 5) (b) Start 2 (Exp. 6)

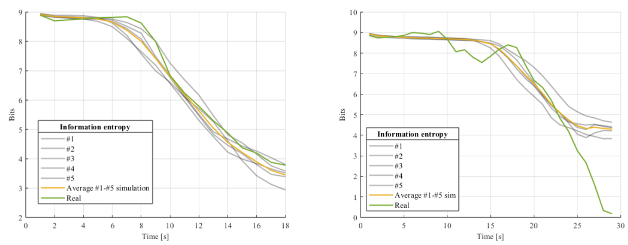
Fig. 5: Source 3: Variance



(a) Real exp. 4 (b) Sim exp. 4

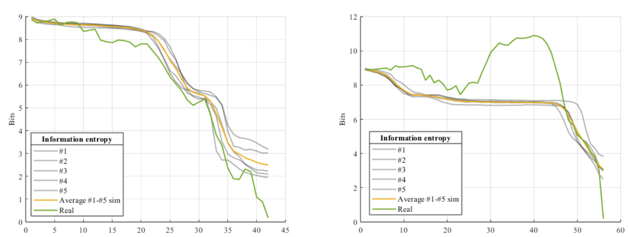
Fig. 6: Source 2: Multivariate kernel density estimation

in information entropy occurs in the last iterations of the algorithm. The real-world algorithm converges more quickly than the simulated one, possibly due to the different behavior of the simulated radiation detector in the surrounding area of the source.



(a) Start 1 (Exp. 1) (b) Start 2 (Exp. 2)

Fig. 7: Source 1: Information entropy

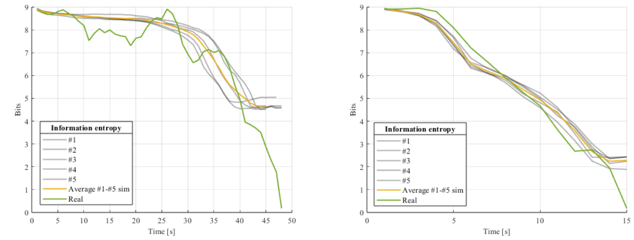


(a) Start 1 (Exp. 3) (b) Start 2 (Exp. 4)

Fig. 8: Source 2: Information entropy

The simulated particle filter requires more measurements near the source for proper convergence. The simulated robot only replicates real-world trajectories, the information entropy does not decrease as rapidly as it does in the real-world algorithm. This aspect is reflected in the Kullback-Leibler divergence. It compares the probability density functions. When the two probability density functions contain identical information, the Kullback-Leibler divergence is zero. Contrarily, when the two distributions differ completely, the divergence goes to infinity.

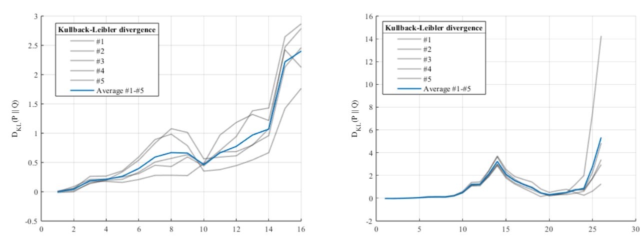
At the beginning of the experiments, the initialization particles represent the same density function because they are



(a) Start 1 (Exp. 5) (b) Start 2 (Exp. 6)

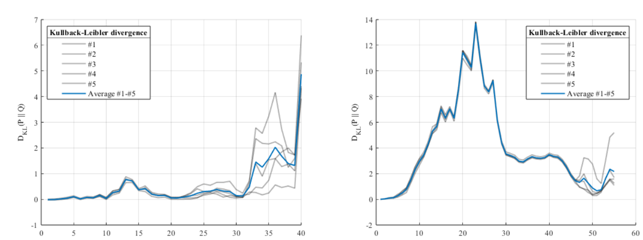
Fig. 9: Source 3: Information entropy

generated from the same distribution. Therefore, the Kullback-Leibler divergence is close to zero, see Fig. 10 - 12. With more measurements, the two particle clouds start to differ slightly, and the divergence gradually increases. The divergence significantly differs in the last measurements before the real-world algorithm ends. The divergence value reaches enormous values due to the fact that the real-world algorithm has already converged while the simulated algorithm has not yet, and the density function completely differs. The particle cloud has a huge density and skewed distribution function at the end of the algorithm. Therefore, the distributions represented by the particles can easily vary a lot. But the estimation can be similar. See example Fig. 13. The final enormous Kullback-Leibler divergence values are not shown in the figures for better resolution.



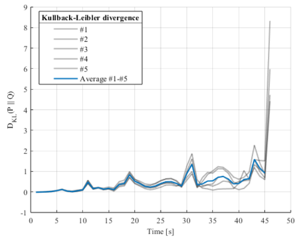
(a) Start 1 (Exp. 1) (b) Start 2 (Exp. 2)

Fig. 10: Source 1: Kullback - Leibler divergence

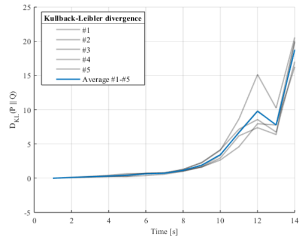


(a) Start 1 (Exp. 3) (b) Start 2 (Exp. 4)

Fig. 11: Source 2: Kullback - Leibler divergence

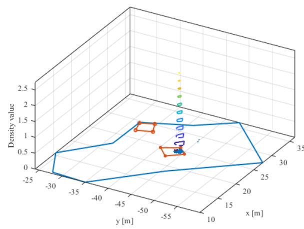


(a) Start 1 (Exp. 5)

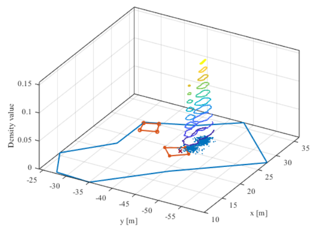


(b) Start 2 (Exp. 6)

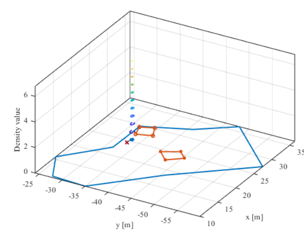
Fig. 12: Source 3: Kullback - Leibler divergence



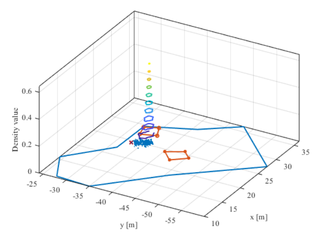
(a) Real experiment 1



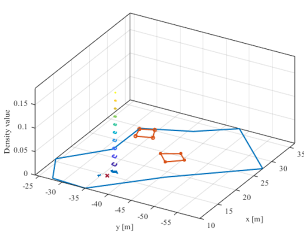
(b) Sim. experiment 1



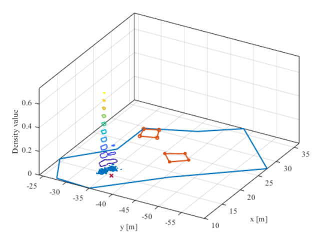
(c) Real experiment 3



(d) Sim. experiment 3



(e) Real experiment 6



(f) Sim. experiment 6

Fig. 13: An example of the probability density function at the end of the algorithm

IV. Conclusion

Radioactivity is a serious threat to both human and animal health and environment. When working with radioactivity in healthcare, industry, and nuclear power plants, it is possible that radioactive sources to become lost, requiring immediate and efficient search to prevent environmental contamination and health hazards. This underscores the importance of locating radioactive sources.

In this paper, we propose and describe a method for searching a radiation source in a specific area using a particle filter with a single sensor. We compared the algorithm's results in both the real world and Gazebo simulator to evaluate the quality of the simulation of the radioactive source and sensor.

By comparing the evolution of particles' information entropy and total variance between the two runs of the particle filter, we have demonstrated the similarities between the simulated and real-world experiments. However, the most significant deviation from reality was observed in the behavior of the radiation sensor near the radiation source. Specifically, the particle filter in the simulation did not converge as quickly as it did in real-world testing, as confirmed by Kullback-Leiber divergence.

The practical implementation of the proposed method was successfully replicated in simulation, indicating that simulation can serve as a valuable tool for the development of mapping and radiation source finding algorithms. By enabling faster and safer testing with sufficient accuracy, simulation has the potential to accelerate future advancements in this field.

Acknowledgment

The research was funded from the Ministry of the Interior of the Czech Republic (MVCR) grant no. VJ02010036 (An Artificial Intelligence-Controlled Robotic System for Intelligence and Reconnaissance Operations).

References

[1] C. East and K. Siewert, "Incident and Trafficking Database".
[2] A. Afzal, D. S. Katz, C. L. Goues, and C. S. Timperley, "A Study on the Challenges of Using Robotics Simulators for Testing." arXiv, Apr. 15, 2020. Accessed: Mar. 11, 2023. [Online]. Available: http://arxiv.org/abs/2004.07368
[3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," IEEE Transactions on Signal Processing, vol. 50, no. 2, pp. 174–188, Feb. 2002, doi: 10.1109/78.978374.
[4] N. Pinkam, A. Elibol, and N. Y. Chong, "Informative Mobile Robot Exploration for Radiation Source Localization with a Particle Filter," in 2020 Fourth IEEE International Conference on Robotic Computing (IRC), Nov. 2020, pp. 107–112. doi: 10.1109/IRC.2020.00024.
[5] P. Xu, C. Fu, Z.-Y. Tan, X. Cai, and J. Qin, "Robust radioactive sources research method using possibility particle filter," AIP Advances, vol. 11, no. 8, p. 085308, Aug. 2021, doi: 10.1063/5.0058860.
[6] S. Thrun, "Probabilistic robotics," Commun. ACM, vol. 45, no. 3, Mar. 2002, doi: 10.1145/504729.504754.
[7] J. Elfring, E. Torta, and R. van de Molengraft, "Particle Filters: A Hands-On Tutorial," Sensors, vol. 21, no. 2, p. 438, Jan. 2021, doi: 10.3390/s21020438.
[8] G. Knoll, Radiation Detection and Measurement. Wiley, 2000.
[9] M. Cihlar, P. Raichl, P. Gabrlik, "Simulation of Autonomous Robotic System for Intelligence and Reconnaissance Operations," Modelling and Simulation for Autonomous Systems (MESAS), 2023
[10] C. E. Shannon, "A mathematical theory of communication," The Bell System Technical Journal, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.
[11] R. Chou, Y. Boers, M. Podt, and M. Geist, "Performance evaluation for particle filters," in 14th International Conference on Information Fusion, Jul. 2011, pp. 1–7.
[12] J. Duchi, "Derivations for Linear Algebra and Optimization".
[13] The MathWorks, Inc. (2022). MATLAB version: 9.13.0 (R2022b). Accessed: January 01, 2023. Available: https://www.mathworks.com

Identifying Anomalies in Industrial Networks: A Proposed Testbed for Experimental Evaluation

Karel Kuchar* and Petr Blazek
Brno University of Technology,
Faculty of Electrical Engineering and Communications, Dept. of Telecommunications,
Technicka 12, 616 00 Brno, Czech Republic
*karel.kuchar@vut.cz

***Abstract*—Not only because of the convergence of Information Technology (IT) and Operational Technology (OT) networks, the pass-through network environment needs to be monitored and adequate security implemented. Due to the occurrence of different types of anomalies and their inconsistency in the literature, three main types of anomalies have been identified in this paper and a testbed has been proposed to serve for further experimental testing. This testbed was created using an anemometer and in the current state using an accelerometer. From the results so far, a correlation between the normal condition and the induced operational anomaly can be observed.**

***Index Terms*—Anemometer, Anomaly detection, Network security, Testbed**

I. Introduction

Industrial Control Systems (ICS) are used to control and automate industrial processes such as manufacturing, production, and more. These systems can be found in a wide range of industries, including energy, transportation, water treatment, and manufacturing. It is thus a combination of hardware resources and software components in order to control and monitor industrial processes. The main objective of ICS is to ensure the efficiency, reliability, process safety, and productivity of industrial processes. Along with the development of industries, another industrial revolution called Industry 4.0 has taken place. There has also been a convergence of Information Technology (IT) and Operational Technology (OT), which has enabled the use of many tools from the IT environment. On the other hand, these technologies have also been exposed to IT threats. Common in OT networks until this time was mainly provided by physical separation, but this has disappeared with the convergence of IT and OT networks and so it is necessary to develop new tools using modern technologies to ensure security in these networks.

To be able to ensure sufficient security level in these types of networks it is needed not only to use new approaches but also to use appropriate methods using all available data. These data might be taken from industrial devices and industrial processes but also from redundant sensors. Redundant sensors might be used to monitor both, industrial processes and devices to get the most possible overview and use correlation evaluate data to evaluate the current state. Current industrial testbeds and dataset [1] mostly focus on incident detection using developed tools mainly via machine-learning and deep neural networks [2]. These approaches are mostly capable of detecting and distinguishing individual attacks using all available data. In other words, it is necessary to provide the maximum amount of information to the incident detection system in order to increase the potential of the tool and to enable not only the correct detection of anomalies but also their classification into defined classes. The data can be divided into main three groups (1) data taken from a transmission medium (protocol data, medium usage, etc.), (2) data taken from the end devices (masters, slaves, servers, actuators, sensors, etc.), and (3) data taken from additional sensors which main purpose is not to control and measure industrial process but to control and measure the state of individual devices.

The purpose of this paper is to present a proposal for the creation of a new industrial testbed that will generate data only from the industrial process but will also use redundant sensors and sensors that will monitor the status of the testbed. The purpose is to further provide proof of work that the use of additional data can be used for early-fault detection and classification of anomalies. Also, the aim of the paper is to introduce the division of anomalies into three categories, namely: security, operational, and service anomaly.

II. Current State-of-the-Art

Due to the heterogeneous approach of anomaly detection and classification in the industrial network, the current state of the art is not consistent from the view of individual anomaly classes. For this reason, we have included in this section relevant resources dealing with the identification of security, operational, and service anomalies. Goran Jurišić, et al focused in their paper [3] on creating of testbed created for analysis of fault detection reaction times. In their approach, they use industrial protocol GOOSE to test the time of reactions. Due to the fact, that the testbed simulates power-grid distribution and impacts of failure, the testbed is also capable of generating different values based on the chosen scenario. In paper [4] Jhonathan Julián Gallego Rojas, et al. introduced real-time Small-Scale Wind Turbine Emulator. This system is used for static behavior analysis of the whole system. The testbed is so not specifically, in this stage, focused on specific anomaly detection. From the view of output data, the testbed is capable to generate diverse data due to the different behavior of the whole system in dependence on wind speed (rotor movement).

Michael Sinner, et al. [5] focused on the development of a model predictive controller for blade pitch control of wind turbines. In their approach, they use as the only one from the selected papers, data from redundant sensors as an additional input of data. They focus on blade pitch control, so the data are not used for anomaly identification/classification. Due to the variable dataset, the output data might vary from the settings of the scenario. Daijiry Narzary and Kalyana Chakravarthy Veluvolu introduced in their paper [6] sensor fault detection method. Their presented method uses data that is normally processed by the system and does not use additional sensors, thus focusing only on early-fault detection. The testbed handles the drive, so the output might be different for each scenario. In paper [7] Jing Wang, et al. described data processing analysis implementation of processing data taken from a penicillin simulator for fault detection. In their approach, they focus on the identification of operational anomalies (fault detection of machines) but this approach is not capable early-fault detection/prediction. The used testbed is able to generate various data.

Paper [8] written by Jinrui Nan, et al. focused on Big data-based early fault detection in relation to batteries. The presented paper describes a method for data extraction for early-fault detection based on big data processed via machine learning approaches. On the other hand, their approach does not detect the existence of attacker and anomaly classification. Paper [9] created by Sinil Mubarak, et al. presents industrial datasets with an ICS testbed. The approach describes a method for data evaluation and anomaly detection presented on industrial protocol Modbus. The data they used do not contain data taken from other external sensors or external machines and their state. Asuka Terai, et al. present in paper [10] cyber-attack detection based on the monitoring system and data evaluation. The approach is presented on a water distribution testbed using industrial protocol OPC. The presented approach is focused only on security incidents. In paper [11] Jonathan Goh, et al. presented a secure water treatment testbed/dataset. The dataset is highly used for data evaluation purposes and detection of cyber anomalies in the industrial system due to the precisely presented dataset consisting of many industrial parts (physical one). The EtherNet/IP protocol is used as a transmission protocol.

Paper [12] described Franck Sicard, et al. developed a physical testbed for naval defense security. In their paper, they describe four possible attacks on that testbed including its detection. In their testbed, they use S7 communication protocol. The resulting dataset contains heterogeneous data due to the heterogeneity of the test environment. Jehn-Ruey Jiang and Yan-Ting Lin presented in their paper [13] anomaly detection technique via deep learning methods. The approach is based on powergrid dataset that uses the Modbus communication protocol. In their paper, they focused only on the detection of security anomalies. Jehn-Ruey Jiang and Yan-Ting Chen presented in their paper [14] anomaly detection and its classification using network traffic. In their paper, they performed the classification of a total of six attacks, i.e. they performed the classification of security anomalies. Their research was based on two publicly available datasets.

Table I provides a summary of the mentioned papers in terms of the parameters assessed. From the table, it is visible that most papers are focused on security anomalies (incidents). Based on the analysis, the approach of evaluating all types of anomalies is not yet used in the current literature. Similarly, the use of additional/redundant sensor data sources and their processing is not frequent. There is also no intersection between early-fault detection and security anomaly detection within industrial networks. There is also no clear definition of the different types of anomalies within the current literature. For this reason, the individual data had to be obtained by classifying the different approaches of the compared papers into defined sections. Values that could not be derived from the content of the paper are marked with a question mark. From the state-of-the-art analysis, it was also found that many papers do not deal with operational anomaly detection and if they do it is typically in combination with early-fault detection. However, many times it is difficult to deduce from the text of the work whether it is possible to perform not only early-fault detection but also anomaly detection at the current time and vice versa.

TABLE I
CURRENT STATE OF THE ART

Paper	Year	RS	Protocol	Anomaly			EFD	Variable data
				Sec.	Oper.	Serv.		
[3]	2018	No	GOOSE	Yes	No	Yes	No	Yes
[4]	2019	No	?	No	No	No	No	Yes
[5]	2021	Yes	?	No	No	No	No	Yes
[6]	2022	No	?	No	Yes	No	Yes	Yes
[7]	2022	No	?	No	Yes	No	No	Yes
[8]	2022	No	?	No	Yes	No	Yes	Yes
[9]	2021	No	Modbus	Yes	No	No	No	Dataset
[10]	2017	No	OPC	Yes	No	No	No	Yes
[11]	2017	No	EtherNet/IP	Yes	No	No	No	Yes
[12]	2022	No	S7	Yes	No	No	No	Yes
[13]	2022	No	Modbus	Yes	No	No	No	Dataset
[14]	2022	No	Modbus/S7	Yes	No	No	No	Dataset

The abbreviation Sec. stands for Security; Oper.: Operational; Serv.: Service; EFD: Early Fault Detection.

III. INDUSTRY 4.0 AND ANOMALY DETECTION APPROACHES

The scope of the fourth industrial revolution called Industry 4.0 is very wide from the perspective of all possible components and features, due to the integration of several digital technologies and concepts in the manufacturing industry, such as the Internet of Things (IoT), artificial intelligence (AI), and big data analytics. Also the convergence of IT and OT technologies opens hitherto completely isolated systems to new threats (from the perspective of OT networks). OT networks were not prepared for this convergence without the effort of additional security features from the perspective of the often-used legacy industrial environment (typically old devices running over an insecure industrial protocol). Therefore, state-of-the-art approaches must be used for the early detection and classification of anomalies within industrial networks. Only through early detection of an intrusion (preferably just an attempted intrusion) can potential damages such as production downtime, data theft/manipulation, equipment damage, safety risk (from the view of the industrial process), data exfiltration (ransomware), malware infiltration (for later usage/backdoor), etc. be avoided. Thus, one of the key aspects appears to be the use of a digital forensic investigation architecture using the maximum amount of data [15].

Anomalies, not only in industrial networks, define a condition where an unusual or unexpected event has occurred within an observed entity (object) that deviates from expected states/values/patterns/behaviour. Based on the purpose and combination of conditions, the anomalies can be divided into three main categories, namely on: security anomaly, operational anomaly, and service anomaly. A security anomaly includes any condition in which a security breach is attempted and can be further divided into subcategories such as Denial of Service (DoS), Man in the Middle (MitM), unauthorized access, and Malware. This category of anomaly can cause targeted network damage, data leakage, denial of service, etc. Thus, it includes all cyber-attack related behavior. Typically this kind of security breach can be detected by analyzing the transmission medium using appropriate intrusion detection and prevention methods and deep analysis such as machine learning techniques or neural networks.

Operational anomalies include all conditions in which an anomaly occurs from the perspective of individual devices as such without the need for targeted intervention or damage. Typically, these can be various types of faults on individual pieces of equipment caused, for example, by material wear or manufacturing defects. The consequence of such anomalies may be deviations in production, and delays in transportation or logistics operations. Operational anomalies can also include early-fault prediction, or the use of data nuances to detect a potential fault early and perform a timely service action without the need to shut down the production process.

Service anomaly refers to all conditions in which an anomaly occurs due to service interventions caused by improper station/element configuration. This can include packet dropping, overloading of some stations, lack of security features, wrong cipher suite, communication errors, slow network connection, excessive network traffic, etc. Some of these events can be detected by collating data obtained from redundant sensors, network traffic and values obtained from individual stations.

IV. Industrial Testbed for Anomaly Classification

In order to perform validation testing of the proposed method, i.e., using redundant sensors to classify the detected anomaly, it is necessary to create an experimental site. Industrial control systems typically work with mechanical devices, their processes or states are monitored or controlled within the process. End devices, i.e. controlled or monitored devices, can be divided into actuators and sensors according to the activity performed. An actuator represents an active device in terms of the operation being performed, its purpose is typically to perform an action (e.g. to close a tap) by means of mechanical work. Sensors, on the other hand, are used to monitor this quantity, i.e., for example, to monitor the flow of a fluid through a monitored pipe by converting the observed phenomenon into an electrical phenomenon (evaluable by software). The resulting correlation, however, can be made using data obtained from the aforementioned actuators and sensors (controlling process) and additional sensors (which are not designed to monitor the observed quantity within the process, but to monitor quantities related to the state of the equipment – for example, vibration). Fig. 1 shows the actuator being controlled through the control station (green line) while data from additional sensors are acquired (orange line). The data is then evaluated at the monitoring and evaluation station using machine learning structures.

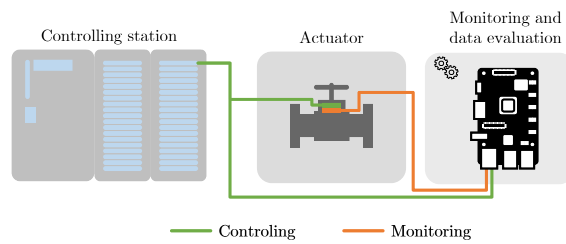


Fig. 1. The principle of advanced data evaluation using redundant sensory sources.

There are certain criteria that need to be taken into account when creating this workplace. The criteria are namely: repeatability, reproducibility, extensibility, based on a real processes, and data variability. The requirement for repeatability focuses on the observed process and the possibility to perform the same experiment repeatedly assuming identical or very similar results. The requirement of reproducibility determines the possibility to create an identical testbed based on the knowledge of the monitored process. The data generated is thus not dependent on a specific workplace, but on the process being monitored. Another requirement is extensibility, which defines the possibility to use other monitored processes or to integrate existing workplaces into another structure, ensuring mutual operability. Furthermore, it is necessary to take into account the requirement for the veracity of the data, or the use of real processes that are monitored and controlled in order to present the possible use in practice and at the same time to minimize the differences from the testbed and real use. Related to this is the requirement for data variability, which defines the variety of data within a process. In other words, it is possible to generate different data for different scenarios, and these data correspond to reality.

A. Testbed description – current state

Based on the established criteria, a testbed was designed. This testbed is based on an anemometer that communicates using the industrial Modbus RTU protocol. This anemometer thus uses the Modbus protocol to acquire individual sensor data such as wind direction and wind speed. It is thus a representation of a sensor that converts the mechanical rotation of the blades relative to the base into an electrical signal, which is then stored in the protocol memory blocks. These register values are then read by the master station where the individual values are stored. Additional data are acquired using a 3-axis accelerometer. Using this data it is possible to identify sensor position, individual vibrations, and acceleration. The Raspberry Pi is connected via a USB cable to the RS-485 converter to achieve Modbus RTU messages. The acceleration sensor (MMA7660FC) is capable to distinguish changes ± 1.5g (1 g = 9.8 m/s^2). The structure is shown in Fig. 2, where the sensor data taken from the anemometer are marked blue and the additional data taken from the additional accelerometer are marked purple. The accelerometer is rigidly attached to the anemometer structure where the induced shocks are acquired. Data is acquired from the accelerometer through the GrovePi+ hat, which is used to connect I2C and other analog ports.

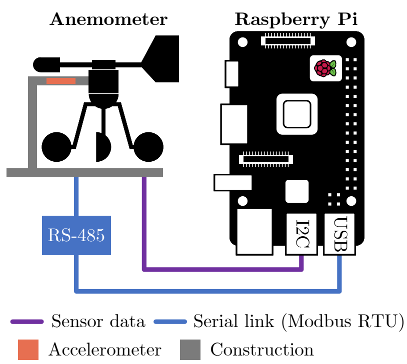


Fig. 2. Mutual interconnection of anemometer and raspberry Pi.

B. Testbed Goals

To create a more purpose-built testbed, further expansion of the workstation with Pulse Width Modulation (PWM) controlled fans is planned. The aim of these fans is to simulate a real environment. In this way, a high-quality dataset covering different states will be obtained and, in particular, this dataset will be diverse in terms of long-term measurements. In addition, other sensors will be placed within the workstation, such as an ultrasonic ranger to measure the vibrations generated by the ultrasonic signal. Subsequently, a data correlation of the measured values from the testbed and the data obtained from the additional sensors is planned. To perform anomaly classification, scenarios focusing on security, operational, and service scenarios will be developed. Based on their analysis, patterns will then be identified and the data will be processed by neural networks to automatically classify the data.

V. Proof of Work

Using the current state of the mentioned testbed, proof of the work is given in this section. With the usage of an acceleration sensor (MMA7660FC) and data taken via Modbus RTU, the two scenarios were described. The first scenario is focused on the legitimate/normal behavior of the testbed. The second scenario is focused on the operational anomaly, so the anomaly behavior is simulated by adding material to one of the rotor blades. The change in wind direction was not considered in this experiment. The wind force was simulated using compressed air. The purpose of this proof of work is only to demonstrate the ability of the basic correlation between the values obtained from the observation of the process (obtaining anemometer readings) and the values obtained from the anemometer. This experiment will be further used to modify the proposed testbed and the use of individual sensors.

Fig. 3 shows the correlation between the individual vibrations as a function of the measured speed from the anemometer in the case of normal operation with no anomalies occurring. In the figure, the speed is shown in red, and the measured vibration is in blue. The average measured velocity was 9.62 m/s, the average vibration rate was 3.24 m/s^2, the median measured velocity was 10.80 m/s and the vibration rate was 2.45 m/s^2. Thus, the ratio of the median values of wind speed and vibration in this case is 2.97 s.

Fig. 4 shows the correlation between the vibration and the measured speed from the anemometer in the case of an operational anomaly. The color coding is identical to Fig. 3. The average measured velocity was 10.85 m/s, the average vibration rate was 6.05 m/s^2, and the median measured velocity and vibration rate were 11.50 m/s and 5.20 m/s^2, respectively. Thus, the median wind speed and vibration ratio, in this case, is 2.21 s. Compared to the normal condition, the ratio was thus reduced by 0.76 s. The operating anomaly was simulated by placing a 2 g load in one of the blades of the anemometer. This is to simulate the clogging of the blade by e.g. a layer of ice. The different scenarios were simulated to achieve the smallest possible differences between scenarios, but deviations still occurred. Therefore, the proposed testbed needs to be refined to ensure data and state consistency.

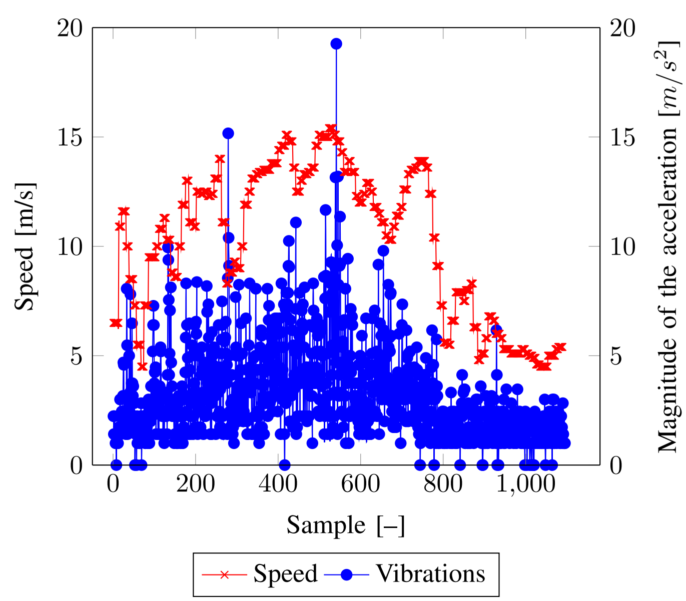


Fig. 3. Values obtained from anemometer and accelerometer - normal behaviour.

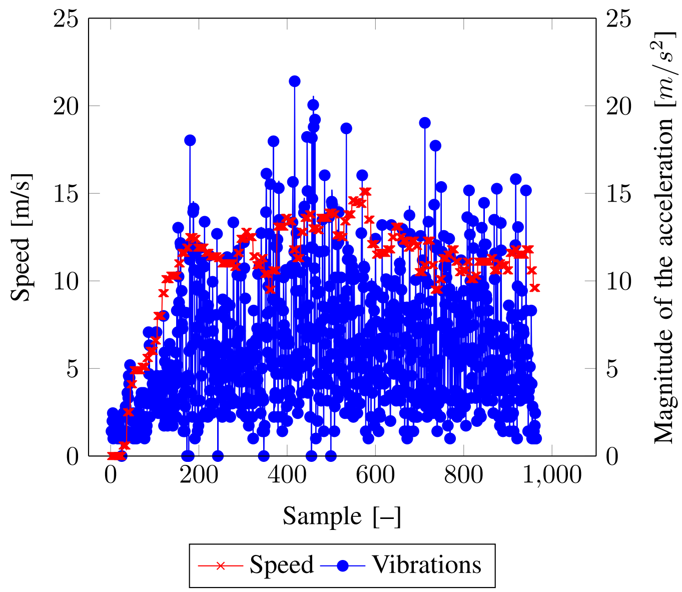


Fig. 4. Values obtained from anemometer and accelerometer - operational anomalies.

VI. Conclusion

This paper focused on presenting three identified classes of anomalies in industrial networks using additional sensor data. Distinguishing these types of anomalies can help in anomaly identification, maintenance planning, and activation of appropriate response measures. A combination of available network data as well as sensor data is needed for appropriate resolution.

Subsequently, an industrial testbed was designed and partially implemented to present these classes of anomalies to obtain a robust dataset. The industrial testbed produced two sensors where the deviations and correlation between measured wind speed and acceleration were observed. The wind speed was obtained through the industrial Modbus RTU protocol and the acceleration was obtained using an accelerometer mounted on the base of the anemometer. The first scenario focused on the correlation during normal running, data was acquired and the ratio of these values was evaluated.

The second scenario simulated an operating anomaly by placing a load in one of the rotor blades presenting, for example, the blade is covered by a layer of ice. From the experimental test, anomalies were identified where the ratio of wind force to total acceleration was 2.97 s in the normal operation and 2.21 s in the second test. Thus, the difference in values is 0.76 s. Thus, by using the designed testbed and the available sensor gauges, it is already possible to differentiate these conditions from each other in this condition and thus identify the operating anomaly. This is despite the limitation that the scenarios were not identical in terms of testing (same wind speed at the same time for both scenarios) and thus cannot be directly compared without incurring additional errors. Thus, the proposed testbed could be used to acquire and evaluate additional data using more types of sensory gauges and to expose additional types and kinds of anomalies.

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 programme.

References

[1] M. Conti, D. Donadel, and F. Turrin, “A survey on industrial control system testbeds and datasets for security research,” vol. 23, no. 4, pp. 2248–2294, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9471765/

[2] S. Olugbade, S. Ojo, A. L. Imoize, J. Isabona, and M. O. Alaba, “A review of artificial intelligence and machine learning for incident detectors in road transport systems,” *Mathematical and Computational Applications*, vol. 27, no. 5, 2022. [Online]. Available: https://www.mdpi.com/2297-8747/27/5/77

[3] G. Jurišić, J. Havelka, T. Capuder, and S. Sučić, “Laboratory test bed for analyzing fault-detection reaction times of protection relays in different substation topologies,” *Energies*, vol. 11, no. 9, 2018. [Online]. Available: https://www.mdpi.com/1996-1073/11/9/2482

[4] J. J. G. Rojas, D. A. Z. Prada, and W. A. L. M. O. Lopez-Santos, “Real-time small-scale wind turbine emulator for a hybrid microgrid laboratory testbed,” in *2019 IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA)*, 2019, pp. 1–6.

[5] M. Sinner, V. Petrović, A. Langidis, L. Neuhaus, M. Hölling, M. Kühn, and L. Y. Pao, “Experimental testing of a preview-enabled model predictive controller for blade pitch control of wind turbines,” *IEEE Transactions on Control Systems Technology*, vol. 30, no. 2, pp. 583–597, 2022.

[6] D. Narzary and K. C. Veluvolu, “Multiple sensor fault detection using index-based method,” *Sensors*, vol. 22, no. 20, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/20/7988

[7] J. Wang, J. Zhou, and X. Chen, *Fault Identification Based on Local Feature Correlation*. Singapore: Springer Singapore, 2022, pp. 119–146. [Online]. Available: https://doi.org/10.1007/978-981-16-8044-1_8

[8] J. Nan, B. Deng, W. Cao, J. Hu, Y. Chang, Y. Cai, and Z. Zhong, “Big data-based early fault warning of batteries combining short-text mining and grey correlation,” *Energies*, vol. 15, no. 15, 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/15/5333

[9] S. Mubarak, M. H. Habaebi, M. R. Islam, A. Balla, M. Tahir, E. A. A. Elsheikh, and F. M. Suliman, “Industrial datasets with ics testbed and attack detection using machine learning techniques,” vol. 31, no. 3, pp. 1345–1360, 2022. [Online]. Available: https://www.techscience.com/iasc/v31n3/44856

[10] A. Terai, S. Abe, S. Kojima, Y. Takano, and I. Koshijima, “Cyber-attack detection for industrial control system monitoring with support vector machine based on communication profile,” pp. 132–138, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7966982/

[11] J. Goh, S. Adepu, K. N. Junejo, and A. Mathur, “A dataset to support research in the design of secure water treatment systems,” *Critical Information Infrastructures Security*, pp. 88–99, 2017. [Online]. Available: http://link.springer.com/10.1007/978-3-319-71368-7_8

[12] F. Sicard, E. Hotellier, and J. Francq, “An industrial control system physical testbed for naval defense cyber-security research,” pp. 413–422, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9799429/

[13] J.-R. Jiang and Y.-T. Lin, “Deep learning anomaly classification using multi-attention residual blocks for industrial control systems,” *Sensors*, vol. 22, no. 23, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/23/9084

[14] J.-R. Jiang and Y.-T. Chen, “Industrial control system anomaly detection and classification based on network traffic,” *IEEE Access*, vol. 10, pp. 41 874–41 888, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9758754/

[15] V. R. Kebande, “Industrial internet of things (iiot) forensics,” *Forensic Science International: Reports*, vol. 5, 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2665910722000032

Test Polygon for the Role of Power Line Communication in ensuring ISO 15118 and IEC 61851 Compliance in EVSE Technology

1st Martin Rusz
Department of Telecommunications
Brno University of Technology, FEEC
Brno, Czech Republic
martin.rusz@vut.cz

2nd Petr Mlýnek
Department of Telecommunications
Brno University of Technology, FEEC
Brno, Czech Republic
mlynek@vut.cz

***Abstract*—The conference paper focuses on the testing of a polygon for the role of Power Line Communication (PLC) in ensuring compliance with ISO 15118 and IEC 61851 standards in Electric Vehicle Supply Equipments (EVSE). The study aims to validate the effectiveness of the proposed polygon as a test platform for evaluating the performance and interoperability of EVSEs that utilize PLC technology for communication with electric vehicle (EV). The proposed polygon is a reliable test platform for evaluating the compliance of EVSEs with ISO 15118 and IEC 61851 standards, and can help in identifying and resolving interoperability issues between different EVSEs and EVs. The paper highlights the importance of ensuring compliance with these standards for the reliable and efficient operation of EVSEs, and demonstrates the potential of the proposed test polygon as a useful tool for achieving this goal.**

***Index Terms*—HLCP, HomePlug Green PHY, IEC 61851, ISO 15118, LLCP, PLC, test polygon**

I. Introduction

As the demand for Electric Vehicles (EVs) continues to increase, there is a growing need for efficient and reliable communication systems between EVs and charging stations (EVSEs). Power Line Communication (PLC) is one such communication technology that is becoming increasingly popular in the electromobility industry. However, before PLC can be deployed on a large scale, it is crucial to test its functionality and reliability in a controlled environment.

This conference paper discuss the importance of a test polygon for testing PLC communication in electromobility. Paper analyze the current state of PLC communication technology, including its advantages and limitations. Moreover, investigate the challenges facing the deployment of PLC communication in the electromobility industry.

Furthermore, this paper will propose the concept of a test polygon for testing PLC communication in electromobility. The article describes the design and implementation of a test polygon, including the selection of appropriate testing equipment, test scenarios, and measurement parameters. Additionally presents the results of testing and provide recommendations for future research and development.

This conference paper aims to provide insights into the importance of testing PLC communication in electromobility and the development of a test polygon for this purpose, similar to [1]–[4]. This paper will contribute to the ongoing discussions on the deployment of efficient and reliable communication systems in the electromobility industry and provide insights for policymakers, industry stakeholders, and researchers.

A. State of the art

The significant increase in the number of electric cars and charging stations increased interest and the need for a bi-directional communication protocol that would enable effective and safe communication between the electric car and the charging station. Current charging stations mainly use the IEC 61851 communication standard, which uses very limited functions for charge management and does not have bi-directional communication.

In recent years, there has been a growing interest in ISO 15118, particularly in Europe. The European Union has recognized the importance of this protocol and has included it in its directive on alternative fuels infrastructure. This directive mandates that all public charging stations in the EU must support ISO 15118 by 2025. For example, in Germany the government has been particularly proactive in promoting the adoption of ISO 15118. In 2016, the German Federal Ministry for Economic Affairs and Energy launched a program called ”Living Lab” to promote the development and testing of ISO 15118-compatible charging infrastructure. The program involved a partnership between the government, automakers, and charging infrastructure providers to accelerate the adoption of ISO 15118.

ISO 15118 enables communication between the electric vehicle and the charging infrastructure, allowing for features such as bi-directional charging, secure communication, and automatic authentication. Overall, the adoption of ISO 15118 is expected to increase as more countries and organizations recognize the importance of this protocol in enabling efficient and secure communication between electric vehicles and charging infrastructure. The protocol’s adoption is also expected to be further accelerated by the increasing popularity

DOI 10.13164/eeict.2023.269

of electric vehicles and the need for interoperability among different charging networks.

B. Contribution

Test polygon for ISO 15118, IEC 61851 and PLC technology can bring significant advantages to the EV charging infrastructure and automotive industries. It can help accelerate the adoption of these protocols, improve interoperability, and ensure the reliability and security of the communication between electric vehicles and charging stations, ultimately leading to a better user experience for EV owners. Polygon will also bring new opportunities in the form of contract research, development of new testing tools and methodologies.

II. Standard for EV charging IEC 61851

IEC 61851 is an international standard that defines a communication protocol for charging electric vehicles. This protocol is used to ensure safe and reliable charging of electric vehicles in various situations and environments. The IEC 61851 protocol defines several types of charging modes. Each of these modes has specific requirements for the equipment used for charging, such as connectors, cables and charging stations [5]. The table I shows the parameters of the modes.

TABLE I
Modes parameters according to IEC 61851-1 [6].

Mode:	Charging Type:	Max. Current:	Max. Power:
1	Slow	16 A, AC, Single-Phase	3.7 kW
2	Fast	32 A, AC, Single-Phase 32 A, AC, Three-Phase	7.4 kW 22 kW
3	Rapid	62 A, AC, Three-Phase	43 kW
4	Ultra-Rapid	400 A, DC	200 kW

- **Mode 1** – this mode does not require any special charging equipment, except for a special cable with protective elements and a connector to connect to the EV.
- **Mode 2** – charging is carried by a special cable with integrated protective elements. This cable contains a communication interface and can be connected to a standard socket or to a special charging station (EVSE).
- **Mode 3** – uses special charging stations with protective elements and a communication interface. Charging in this mode is secured by introducing standardized charging connectors and an interface for communication between the EV and the EVSE.
- **Mode 4** – mode for DC charging at a permanently installed (EVSE). The EV can be charged using two different connector systems: Combined Charging System (CCS) and CHAdeMO.

The IEC 61851 also specifies various parameters for charging, such as voltage, current, frequency and protection elements, which ensure the safety and reliability of charging. In addition, it defines communication protocols for communication between the EVSE and the EV, which includes vehicle identification, charging status information, charging control options, and more.

A. Principles of communication within IEC 61851-1

Communication within the IEC 61851-1 standard between EVSE and EV is based on pulse width modulation (PWM) between Control Pilot (CP) and Ground (GND) wires. The main goal is to ensure the transmission of all the necessary energy related information, so called Low Level Communication Protocol (LLCP). The figure 1 shows the individual states/faults and the corresponding value of the PWM signal.

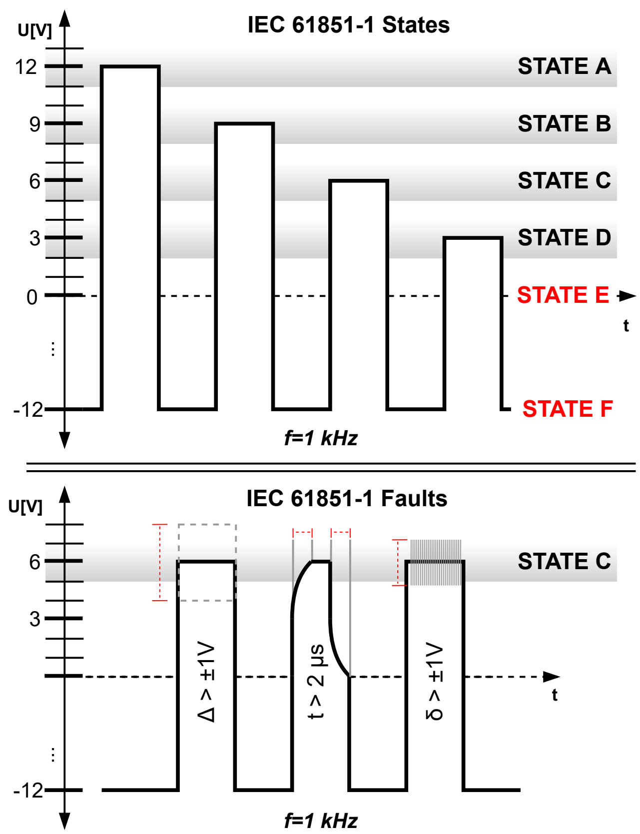


Fig. 1. Indication of states and faults according to the PWM signal [7].

Voltage +12 V on the CP wire signals state A, which means that the EV is not connected. After connecting the EV to the EVSE, the positive value of the PWM signal amplitude will be reduced to +9 V through a simple circuit, and state B is signaled (EV connected). The start of charging is requested by the EV by connecting an additional resistor to the circuit, the voltage is reduced to +6 V and status C (charging) is signaled. If additional ventilation is needed, an additional resistor is connected and the voltage value of the PWM signal drops to +3 V, state D is indicated (ventilation required). In states A–D, the negative part of the PWM signal must be -12 V. State E indicates a problem on the EVSE side and state F indicates that the EVSE is not available. Reliability of communication even during interference is ensured by the permitted deviation for the voltage level and distortion of the PWM signal.

The second parameter that can be transmitted as part of the communication is the maximum value of the current that the EV can take from the EVSE or grid. The information about the maximum value of the current is transmitted by duty cycle of the PWM signal. The transmission of this information is critically important to ensure the stability and security of the grid by preventing overload. The circuit for ensuring IEC 61851 communication between EV and EVSE is shown in the figure 3.

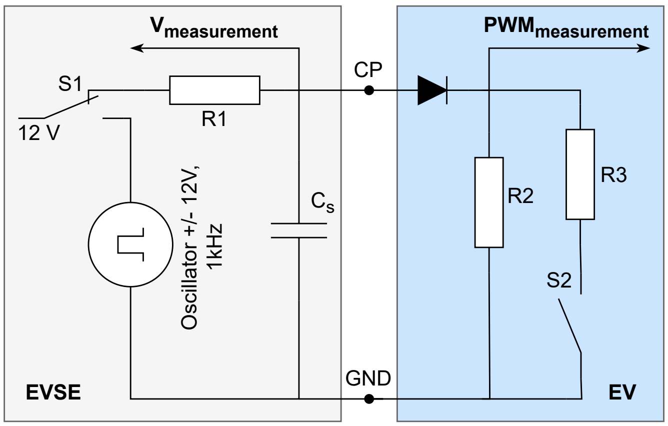


Fig. 3. Simplified circuit diagram for communication between EVSE and EV.

III. Communication protocol ISO 15118

ISO 15118 is a communication protocol standard that facilitates the communication between EV and EVSE. It specifies the information exchange requirements, communication sequences, and messages for bi-directional communication. Key component of the ISO 15118 standard is the High-Level Communication Controller (HLLC), which allows the use of High Level Communication Protocol (HLCP) [8]. The HLLC serves as the interface between the EV and the EVSE. The HLLC is responsible for managing the communication between the EV and the EVSE, and ensuring that the communication is secure, reliable, and efficient. The HLLC performs several key functions:

- **Authentication** – HLLC ensure that EV and EVSE are authorized to communicate with each other. This is important for security reasons, as it prevents unauthorized devices from accessing the EVSE.
- **Identification** – verifying that EV and EVSE are compatible with each other. This ensures that the EV and the EVSE can communicate effectively and efficiently.
- **Session establishment** – is responsible for establishing a secure communication session between the EV and the EVSE. This involves negotiating the communication parameters, such as the communication speed and the encryption algorithm.
- **Message exchange** – is responsible for managing the exchange of messages between the EV and the EVSE. This includes sending and receiving messages, as well as managing message errors and retries.
- **Security** – ensuring that the communication between the EV and the EVSE is secure. This involves encrypting the communication and using authentication mechanisms to prevent unauthorized access.

A. *Implementation of High Level Communication*

The figure 2 shows the principle of communication between EV and EVSE within the ISO 15118 protocol using PLC technology, which is described in the chapter IV. The existing communication according to IEC 61851 is further extended by PLC modems of the HomePlug Green PHY (HPGP) standard. Modems implement the necessary functions to ensure bi-directional secure communication, according to the ISO 15118 protocol.

HPGP Modems are specifically designed for emobility applications and have the SLAC (Signal Level Attenuation Characterization) protocol, which is designed to meet the requirements of the ISO 15118 protocol. SLAC is a protocol that enables the EV and the EVSE to establish a secure communication channel and exchange authentication and authorization information. SLAC protocol involves several steps:

- **Synchronization** – synchronization process ensures that both the EV and the EVSE are synchronized to the same frequency and time.
- **Network Discovery** – EV and the EVSE start the network discovery process to discover each other on the

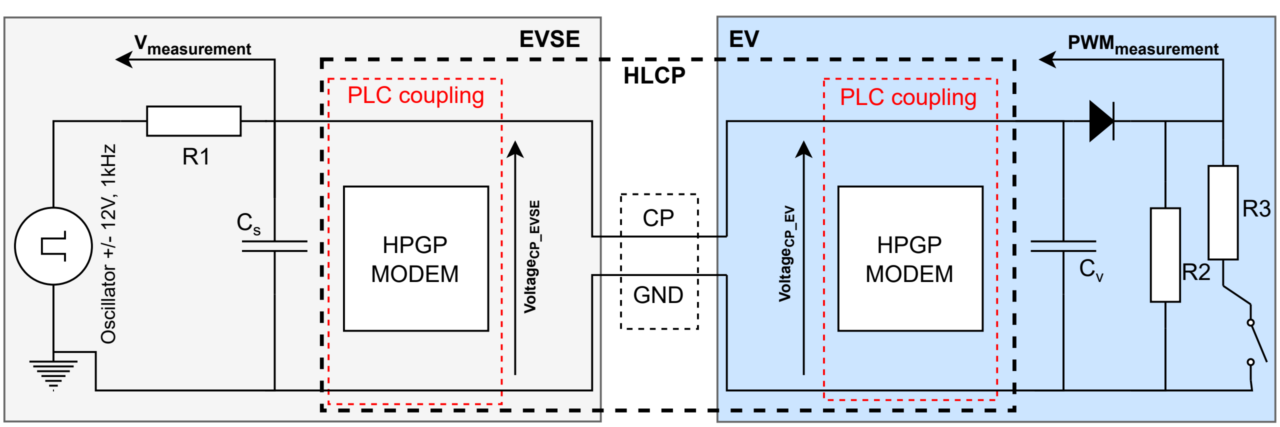


Fig. 2. HLCP implementation between EVSE and EV [9].

power line network. The EV sends a broadcast message to discover available EVSEs, and the EVSE responds with a unicast message.

- **Link Setup** – process establishes a communication link between the EV and the EVSE. The link setup process involves several steps, including signal quality estimation, link negotiation, and key exchange.
- **Authentication** – ensures that the EV and the EVSE are authorized to communicate with each other. The authentication process uses advanced encryption algorithms to protect the transmitted data from unauthorized access.

IV. Role of the PLC in emobility industry

In the context of ISO 15118, PLC is used to establish a communication channel between the EV and the charging station. This communication channel is used for a range of functions, including authentication and authorization, data exchange, and control of the charging process.

One of the key benefits of using PLC for EV charging is that it eliminates the need for additional wiring or infrastructure. This makes it a cost-effective solution that can be easily implemented in a range of different environments. It also provides a reliable and secure communication channel, which is crucial for ensuring the safety and efficiency of the charging process.

A. HomePlug Green PHY

HomePlug Green PHY (HPGP) is a communication technology that is used for powerline communication (PLC) in various applications, including the electric vehicle (EV) charging infrastructure. One of the key advantages of HomePlug Green PHY is its energy efficiency. The technology includes power management features that enable the EV charging process to be more energy-efficient. This is particularly important during peak demand periods when the electricity grid is under strain. HomePlug Green PHY helps to minimize the impact of EV charging on the grid by reducing energy consumption during the charging process [10].

The HomePlug Green PHY technology is also highly adaptable, making it an ideal solution for the evolving EV charging infrastructure. The technology can be used in conjunction with other communication technologies, including cellular communication and Ethernet, to provide a seamless and secure communication channel between the EV and the charging station.

HPGP is directly based on the previous HomePlug AV (HPAV) standard, which was primarily intended for broadband transmission of audio and video content. HPGP share several similarities with in terms of their physical layer technology, interference mitigation techniques, security mechanisms, backward compatibility, standards compliance, and energy efficiency. HPGP focuses mainly on reliable and energy-efficient communication instead of high data rate. The table IV-A describes the differences between HPGP and HPAV standards.

TABLE II
Comparison of the HomePlug AV standards.

Parameter:	HP AV	HP GP
Freq. Band	2–30 MHz	2–30 MHz
Modulation	OFDM	OFDM
Number of subcarrier	1155	1155
Subcarrier modulation	BPSK, QPSK, 16–1024 QAM	QPSK
Supported data transfer	ROBO (up to 10 Mbit/s) ABL (up to 200 Mbit/s)	ROBO

V. Concept of the test polygon

The designed test polygon, shown in the figure 5, presented in this article, provides a controlled environment for testing the implementation of IEC 61851 and ISO 15118 protocols within EVSE of various manufacturers. It is designed to simulate

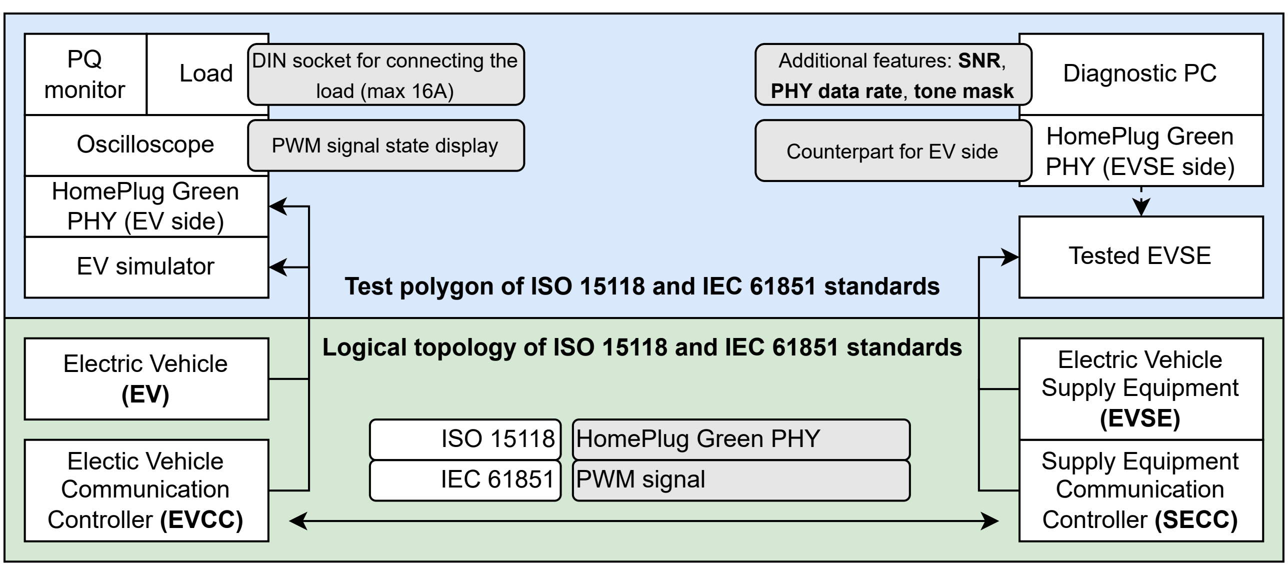


Fig. 4. Block diagram of the proposed test polygon.

conditions and test the performance and reliability of EV charging systems under different scenarios.

The test polygon is designed to exactly match the logical topology of EVSE – EV communication and IEC 61851 and ISO 15118 standards. The part representing the EV is made up of several sub-units. To simulate the individual states according to IEC 61851, described in the subsection II-A, a simple analog circuit was designed, which simulates the individual EV states using a dip switch and several resistors. An oscilloscope is connected to the analog circuit simulating EV to measure the voltage level of individual states, their visualization and communication diagnostics. The simulation of energy consumption is realized by a DIN socket connected to the EVSE. The polygon also include a PQ (Power Quality) monitor for control measurement of current drawn by load connected to socket.

Polygon also enables communication testing according to the ISO 15118 protocol. For this purpose, two HPGP modems are deployed in the polygon, which are capable of simulating the entire communication between EVSE and EV or testing real EVSE with the use of EV side HPGP. Modems implement High Level Communication (HLC), which was described in more detail in subsection III-A. Using the Open-PLC-Toolkit, the diagnostic PC (Raspberry Pi) also enables monitoring of parameters such as SNR (Signal to Noise Ratio), data rate on physical layer or tone mask. The assembled polygon shows the figure 5.

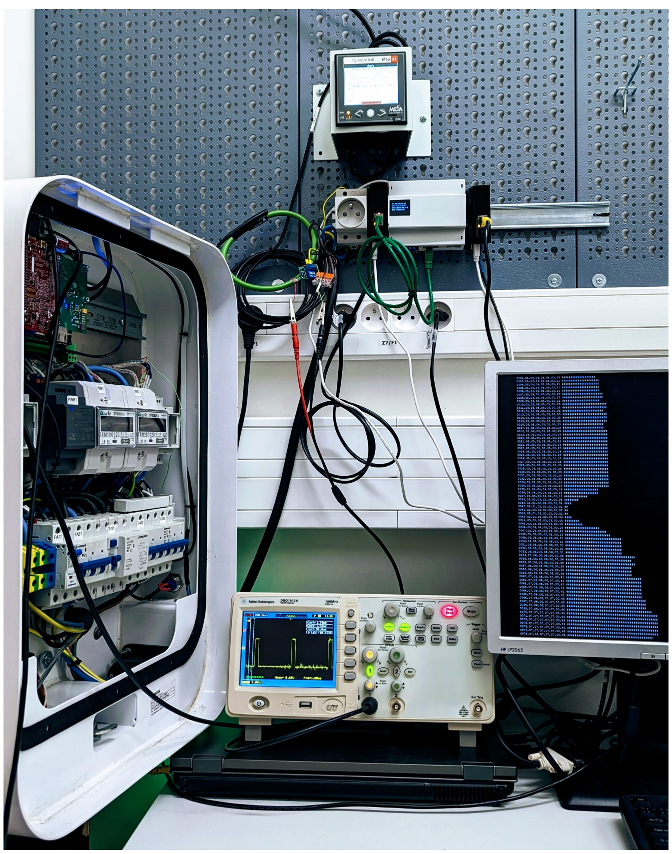


Fig. 5. Assembled polygon for EVSE testing.

In the figure we can see a real EVSE measurement with implemented IEC 61851 communication. The oscilloscope displays the current charging status and ccurrent draw defined by duty cycle. The constructed polygon will be further expanded with other elements such as alternative HPGP modems, advanced EV simulator and SW tools.

VI. Conclusion

The article describes the design of a polygon for testing the IEC 61851 and ISO 15118 standards for communication between EVs and EVSEs. In the individual chapters, the IEC 61851 and ISO 15118 protocol are described in detail. Part of these chapters is also the role of PLC communication, specifically HomePlug Green PHY and its application within the mentioned ISO 15118 communication protocol. The designed polygon is first presented on a block diagram and the individual parts of the polygon and their meaning and function are described. Subsequently, an image of the constructed polygon during real EVSE testing is shown. Polygon is currently used for testing several charging stations with the implemented IEC 61851 protocol.

Future work will focus on the design of test scenarios, collection and evaluation of measured data, design of measurement methodologies and possible HW and SW expansion of the polygon.

This work was supported by the Technology Agency of Czech Republic project No. TK05020114, Czech Republic.

References

[1] M. Shin, H. Kim, H. Jang, "Building an Interoperability Test System for Electric Vehicle Chargers Based on ISO/IEC 15118 and IEC 61850 Standards," Applied Sciences, 2016, doi: 10.3390/app6060165.

[2] Lathika and S. Nithin, "Development of a Communication Test Bed for Electric Vehicle Charging," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 1127-1130, doi: 10.1109/ICESC48915.2020.9155929.

[3] E. Ucer and M. J. Kisacikoglu, "Design and Implementation of a Hardware Test-bed for Real-time EV-Grid Integration Analysis," 2022 IEEE Energy Conversion Congress and Exposition (ECCE), Detroit, USA, 2022, pp. 1-8, doi: 10.1109/ECCE50734.2022.9947704.

[4] Z. Wu, O. Alizadeh, M. Lelic, W. Irfan and R. Gerdes, "Test Bed Development for Evaluating Extremely Fast Charging Stations," 2021 IEEE Transportation Electrification Conference & Expo (ITEC), Chicago, USA, 2021, pp. 2-8, doi: 10.1109/ITEC51675.2021.9490122.

[5] A. Di Giorgio, F. Liberati and S. Canale, "IEC 61851 compliant electric vehicle charging control in smartgrids," 21st Mediterranean Conference on Control and Automation, Platanias, Greece, 2013, pp. 1329-1335, doi: 10.1109/MED.2013.6608892.

[6] A. Triviño, J. Gonzalez, J. Aguado, "Wireless Power Transfer Technologies Applied to Electric Vehicles: A Review" Energies, (2021), volume 14, doi: 10.3390/en14061547.

[7] S. Gröning, C. Lewandowski, J. Schmutzler and C. Wietfeld, "Interoperability Testing Based on TTCN-3 for V2G Communication Interfaces," 2012 International Conference on Connected Vehicles and Expo (ICCVE), Beijing, China, 2012, pp. 298-303, doi: 10.1109/ICCVE.2012.67.

[8] M. Mültin, "ISO 15118 as the Enabler of Vehicle-to-Grid Applications," 2018 International Conference of Electrical and Electronic Technologies for Automotive, Milan, Italy, 2018, pp. 1-6, doi: 10.23919/EETA.2018.8493213.

[9] International Standard, "Road vehicles — Vehicle to grid communication interface — Part 3: Physical and data link layer requirements", ISO 15118-3, first edition, 2015.

[10] Standard whitepaper, "The Standard for In-Home Smart Grid Powerline Communications: An application and technology overview", HomePlug Alliance, 2012.

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