Informatics Control Measurement
in Economy and Environment Protection

1/2021

Informatyka Automatyka Pomiary

w gospodarce i ochronie środowiska

www.e-IAPGOS.pl

ISSN 2083-0157

Kwartalnik Naukowo-Techniczny

fot. Stanislava Moroza

V. M. Glushkov Institute of Cybernetics
of the National Academy of Sciences of Ukraine
(Kyiv, Ukraine)

1/2021

styczeń – marzec

Wydanie pod redakcją naukową
prof. dr hab. inż. Waldemara Wójcika

INFORMATYKA AUTOMATYKA POMIARY

W GOSPODARCE i OCHRONIE ŚRODOWISKA

Informatics Control Measurement in Economy and Environment Protection

p-ISSN 2083-0157, e-ISSN 2391-6761, www.e-iapgos.pl

PRINTING HOUSE – DRUKARNIA

DjaF – Naświetlarnia B1+
ul. Kmietowicza 1/1
30-092 Kraków
http://www.djaf.pl
nakład: 100 egzemplarzy

OTHER INFORMATION – INNE INFORMACJE

Czasopismo jest indeksowane w bazach:

DOAJ:	doaj.org
BazTech:	baztech.icm.edu.pl
IC Journals Master List:	www.journals.indexcopernicus.com
Google Scholar	scholar.google.pl
POL-index	pbn.nauka.gov.pl
Sherpa RoMEO	www.sherpa.ac.uk

Czasopismo *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska* zostało objęte finansowaniem przez Ministerstwo Nauki i Szkolnictwa Wyższego w ramach programu *Wsparcie dla czasopism naukowych* w latach 2019-2020.

Czasopismo znajduje się w wykazie czasopism naukowych opublikowanym w Komunikacie Ministra Edukacji i Nauki z dnia 9 lutego 2021 r., Unikatowy Identyfikator Czasopisma: 200167 – z przypisaną liczbą punktów przyznawanych za publikację artykułu równą 20.

Zasady publikowania artykułów, przygotowania tekstów, zasady etyczne, procedura recenzowania, wykazy recenzentów oraz pełne teksty artykułów dostępne są na stronie internetowej czasopisma:

www.e-iapgos.pl

W celu zwiększenia oddziaływania czasopisma w środowisku naukowym redakcja zaleca:

- w artykułach publikowanych w IAPGOS cytować artykuły z renomowanych czasopism międzynarodowych (szczególnie indeksowanych w bazach Web of Science oraz Scopus) używając oficjalnych skrótów nazw czasopism,
- w artykułach publikowanych w innych czasopismach (zwłaszcza indeksowanych w bazach Web of Science oraz Scopus) cytować prace publikowane w IAPGOS – zwłaszcza posługując się numerami DOI, np.:
  Kluszczyński K. *Modelowanie – umiejętność czy sztuka?* Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOS, 1/2016, 4–15, DOI: 10.5604/20830157.1193833.

CONTENTS – SPIS TREŚCI

http://doi.org/10.35784/iapgos.2456

THE SYSTEM FOR COMPLEX MAGNETIC SUSCEPTIBILITY MEASUREMENT OF NANOPARTICLES WITH 3D PRINTED CARCASS FOR INTEGRATED RECEIVE COILS

Mateusz Midura, Przemysław Wróblewski, Damian Wanta, Grzegorz Domański, Mateusz Stosio, Jacek Kryszyn, Waldemar T. Smolik

Warsaw University of Technology, Faculty of Electronics and Information Technology, Institute of Radioelectronics and Multimedia Technology, Division of Medical and Nuclear Electronics, Warsaw, Poland

Abstract. *The article concerns the research on the properties of core-shell superparamagnetic nanoparticles in the context of their use in medicine for diagnostics and therapy. The article presents a system for impedance (AC) spectroscopy of nanoparticles with a new arrangement of receive coils. A significant modification was the position of the reference coil in relation to the receive coils as well as the method of winding and routing the wires on the carcass. The 3D printing technique was used in the production of the measuring coil system. The aim of the work was to experimentally verify the developed measurement system and analyze its properties. The system tests were carried out at low frequencies ranging from 2 to 50 kHz. Complex magnetic susceptibility was measured for superparamagnetic iron oxide nanoparticles in polymer shells in a physiological saline solution. The obtained results confirmed the relevance of the concept of the measurements. In summary, the observed properties of the realized system are discussed and further directions of its development are proposed.*

Keywords: superparamagnetic nanoparticles, magnetic particle spectroscopy, magnetic susceptibility, hyperthermia

SYSTEM DO POMIARU ZESPOLONEJ PODATNOŚCI MAGNETYCZNEJ NANOCZĄSTEK Z WYKONANYM W TECHNOLOGII DRUKU 3D KARKASEM ZINTEGROWANYCH CEWEK ODBIORCZYCH

Streszczenie. *Artykuł dotyczy badań właściwości nanocząstek superparamagnetycznych typu rdzeń-powłoka w kontekście wykorzystania ich w medycynie do diagnostyki jak i terapii. W artykule przedstawiono układ do spektroskopii impedancyjnej (AC) nanocząstek z nowym układem cewek odbiorczych. Istotną modyfikacją była pozycja cewki referencyjnej względem cewek odbiorczych jak również sposób nawijania i prowadzenia przewodów na karkasie. W realizacji układu cewek pomiarowych wykorzystana została technika druku 3D. Celem pracy była eksperymentalna weryfikacja opracowanego układu pomiarowego i analiza jego własności. Testy układu zostały przeprowadzone dla niskich częstotliwości w zakresie od 2 do 50 kHz. Pomiary zespolonej podatności magnetycznej dokonano dla nanocząstek superparamagnetycznych tlenku żelaza w otoczkach polimerowych w roztworze soli fizjologicznej. Uzyskane wyniki potwierdziły poprawność koncepcji realizacji pomiarów. W podsumowaniu omówiono zaobserwowane własności zrealizowanego układu i zaproponowano dalsze kierunki jego rozwoju.*

Słowa kluczowe: nanocząstki superparamagnetyczne, spektroskopia cząstek magnetycznych, podatność magnetyczna, hipertermia

Introduction

The extensive application of magnetic nanoparticles has attracted a great attention to their measurable properties. The potential wide application of magnetic nanoparticles has attracted a lot of attention from researchers representing various fields of science and has resulted in great interest in measurement methods that enable their characterization.

In medicine, magnetic nanoparticles are applied in diagnostics as well as in therapy, e.g. as a tracer in magnetic particle imaging (MPI) or carriers in drug delivery and in microfluidic devices [8] or as a heating medium in hyperthermia.

Magnetic nanoparticles may have different morphology but typically, those used in hyperthermia, are composed of superparamagnetic iron core and biocompatible coating (shell). The shell layer around a functional nanoparticle core provides a possibility to modify their physical and chemical properties, and allows binding of specific ligands such as drug molecules. One of the magnetic properties of the core that characterises nanoparticles is complex magnetic susceptibility. Imaginary part of susceptibility corresponds to magnetic losses generated due to alternating magnetic field.

In magnetic hyperthermia [4, 16] as well as in magnetic particle imaging it is crucial to determine the optimal frequency at which extreme values of power loss are obtained. In the first application, magnetic losses of nanoparticles should be maximized in the tumour volume, whereas at the same time losses generated due to Eddy currents in patient's body should be as low as possible. In MPI any magnetic losses causing undesired heating of the body should be reduced to minimum.

Significant research problem is to model nanoparticle properties and develop measurement method dedicated to specific applications. In magnetic hyperthermia the magnetic loss is critical parameter directly connected with an ability to generate heat. It strictly depends on hydrodynamic diameter (Dh), which is difficult to measure directly. The only method for direct Dh evaluation is Dynamic Light Scattering (DLS) [9]. The geometric sizes of the produced nanoparticles vary in the range from 5 nm to 100 nm (core diameter). Frequently, the value declared by the manufacturers is approximate and does not include particle's diameter dispersion.

Although several techniques for the synthesis of metal-based core-shell nanoparticles are known, the production of MNPs with well controlled magnetic parameters remains a challenge. A knowledge of the accurate hydrodynamic diameter is necessary for correct modelling of magnetization. It is well known that following factors affect the magnetic susceptibility: the hydrodynamic diameter, the diameter of the metallic core and biocompatible coating, the shape and the material of the core and the shell, as well as the medium in which the nanoparticles are dispersed.

Magnetic properties of nanoparticles can be measured in an immobilize state or as a liquid suspension. These properties differ greatly depending on the environment which determines the degree of mobility. Although, results for immobilized particles are highly dependent on immobilization procedure, they have been quite well documented [1]. Whereas the measurement of magnetic particle suspension still needs standardization and better models for interpretation of the results. There is a need to develop robust method for measuring complex magnetic susceptibility of MNPs which is critical in assessing the effectivity and safety in medical applications.

Magnetic Particle Spectroscopy (MPS) is commonly used for determination of nanoparticle properties like hydrodynamic diameter distribution or magnetic core diameter [14]. These parameters are necessary for calculating complex magnetization of an investigated sample. However, there are also another methods such as Nuclear Magnetic Resonance (NMR) [15]

or Transmission Electron Microscopy (TEM) that allow to evaluate nanoparticle properties [9]. MPS was successfully applied to verify theoretical models of magnetization [3, 7].

A commonly used devices to measure magnetic properties of nanoparticles are: Vibrating-Sample Magnetometer (VSM) [11], Magnetic Particle Spectroscope (MPS) [2, 10, 18], Super-paramagnetic quantifier (SpaQ) [6] and AC magnetometer [13].

In VSM method higher magnetic fields cause the saturation of examined nanoparticles and thereby their magnetic moment remains constant. It is also a well-known fact that VSM measurements at low fields cause problems of magnetic remanence in the system when superconducting coils (SQUID) are used [20].

In a SPaQ a small AC field is combined with a gradual DC offset. It analyses only the fundamental frequency of the detection signal whereas MPS measures the whole harmonic spectrum caused by the nonlinearity of the magnetization. MPS can be described as a dynamic measurement of the magnetization curve compared to the VSM or SPaQ. Both the SpaQ and MPS can be conducted at room temperature, provide accurate measurements of the magnetization curve, and both are superior to VSM especially at low amplitudes of magnetisation field [6].

Whereas all these methods use typically a single excitation frequency, AC magnetometer (impedance spectroscopy) applies multitude of frequencies to obtain spectral magnetic response of a sample. Most bench-top systems are able to analyse only small sample quantities due to the requirements of strong magnetic field. A miniaturized measurement system for AC magnetometry was also developed as it was presented in a project by Šouc et al. [12]. The effort is made to improve sensitivity, signal purity and measurement accuracy.

In this paper, we verify a laboratory setup for measurement of complex magnetic susceptibility by AC impedance spectroscopy method. For this purpose, the system for measuring complex magnetic susceptibility in function of frequency was created. The novelty of the elaborated experimental setup lies in a symmetrical placement of a reference coil positioned in equal distance between two detection coils. The applied 3D-printing technology allowed to fabricate custom-designed carcass for proposed arrangement of coils and ensured proper fitting of carcass in the excitation coil.

1. Methodology

1.1. System configuration

Laboratory setup for impedance spectroscopy is composed of a transmit circuit for generating the sinusoidal magnetic excitation field and a receiving circuit for detecting the MNP response in the form of non-sinusoidal magnetization. To obtain high value of an exciting magnetic field, a resonant circuit is used in the transmitting part of the system. The receiving circuit requires an application of a cancellation coil in pair with a detection coil to cancel the excitation signal in the receiving circuit. This way only the magnetization signal from MNPs located inside receiving coils is present in the detecting circuit. The amplitude and the phase of magnetization signal is measured using a phase-sensitive detector.

Schematic chart of the system used in measurements is shown in Fig. 1. Self-designed excitation coil is a part of resonant circuit. Different sets of capacitors were used to tune resonant frequency. Function generator connected to power amplifier was used as a signal source. Two receiving coils have been wound in the opposite directions and placed inside excitation coil. This configuration allows to cancel excitation signal generated due to the excitation field in receiving coils and was widely discussed in [12, 13, 19]. Therefore, all harmonic components of magnetization signal can be measured by the lock-in amplifier. Lock-in amplifier or phase-sensitive detector is quite a versatile measurement instrument for extracting low-level signals buried in noise [5].

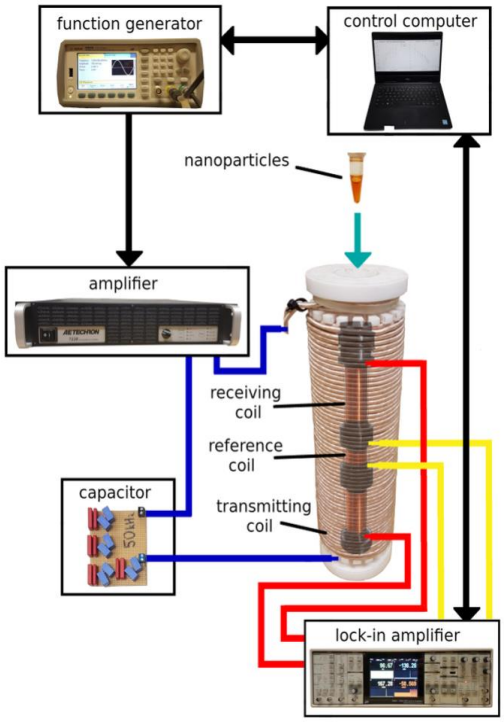


Fig. 1. Scheme of top-bench setup for impedance spectroscopy

1.2. Excitation coil

The transmit coil is a solenoid coil (length: 200 mm, inner diameter: 30 mm, outer diameter 60 mm, overall turns 220) (Fig. 2a). The bore in the carcass of excitation coil was designed to enable inserting the receiving coils together with the measured sample. Detailed design of excitation coil with the analysis of magnetic field distribution was presented in [17]. Litz wire was used in order to minimize skin effect, which cause reduction of current in excitation circuit for higher frequencies. The litz wire contains 7 bundles and each bundle is composed of 630 strands of 0.1 mm copper wire. To protect carcass from melting due to heat generated by excitation coil PTFE (Teflon) was used as a construction material. Active cooling of transmit coil was not necessary during the experiment. Electric parameters of designed coil are presented in Tab. 1. The skin effect was inevitable for frequencies higher than 30 kHz. Thus to maintain the constant amplitude of excitation magnetic field it was necessary to increase gain of the amplifier accordingly.

Table 1. Electric parameters of transmit coil

L [μH]	R (<25 kHz) [Ω]	Q (20 kHz)	Z (<1 kHz) [Ω]	N	L [mm]
747.95	0.22	427	<5	220	200

Fig. 2. Coils used in the measurement setup: a) transmit coil, b) receive coils

1.3. Receive and reference coils

The receive coil system (Fig. 2b) is composed of two opposite wound serial coils: detection coil, cancelation coil and reference coil. Electric parameters of these coils are shown in Tab. 2.

Table 1. Electric parameters of receive coils

Coil	L [μH]	R (<25 kHz) [Ω]	Q (20 kHz)	Z (<1 kHz) [Ω]	N	l [mm]
Receive	139.66	2.88	3.0	3.01	198	128.5
Detection	74.3	1.44	3.2	1.52	99	34
Cancellation	69.25	1.44	3.0	1.51	99	37
Reference	13.60	0.6	1.6	0.6	28	10

The detection coil serves for measurement of magnetization signal of nanoparticles whereas the main purpose of cancelation coil is to eliminate the induced signal by primary magnetic field.

Cancelation is achieved when signal from emptied receive coils is brought to minimum. This can be obtained by positioning the carcass inside the excitation coil using a linear actuator.

The reference coil is integrated with receive coils. That means it is placed inside an excitation coil's bore together with the receive coils on the same carcass. The reference coil measures excitation field signal that provide a frame of reference to lock-in amplifier. It also allows to estimate the strength of magnetic field, that can be used in feedback loop to control the amplitude of function generator signal.

In our design a significant modification was the position of the reference coil in relation to the receiving coils as well as the method of winding and routing the wires on the carcass. The 3D printing technique was used in the implementation of the measuring coil system.

Detection coils' support was designed and built using 3D-printer. Both receive coils and reference coil were wound around 3D-printed carcass. The isolated copper wire of 0.6 mm was used Reference coil was placed in the gap between receiving. Induced voltage was used as a reference signal for lock-in detector.

The design of carcass was prepared using Autodesk Inventor Professional 2021 (Fig. 3). To minimize the mutual inductance between receiving and cancelling coils, both coils' supporting shafts with a length of 34.50 mm has been pulled apart at a distance of 59.50 mm. The outer diameter of 29 mm has been chosen to fit bore diameter of excitation coil (30 mm). Arrangement of reference coil in the gap between detection coils required design of additional holes that served for the proper guiding of both reference coil and detection coils wires. Both ends of each cooper wire were twisted along the length of the cable to eliminate common interference induced by the primary excitation field. Two cable ties were used to ensure a firm, secure soldering and attachment of coaxial cable to 0.6 mm copper wires. For this purpose two troughs were designed to accommodate the flexible tape and the head of cable ties. The upper inlet was adjusted for the size of test tube's cup. It allowed a repeatable stable placement of samples during experiment. At the bottom of carcass the hole was left for inserting a manually controlled linear actuator that was used for calibration.

Using measurement setup with novel arrangement of reference coil for given frequency we were able to maintain constant phase shift between reference and detection coil. Using 3D-printing method for carcass construction significant resemblance in geometric properties of both detection and cancelling coil has been achieved. It allowed to reduce the unbalanced sinusoidal signal induced in receiving coils to approximately 5 mV. As we expected, the placement of reference coil did not impair cancellation of excitation signal in receiving coils.

The designed carcass was printed on Ultimaker S3 3D-printer with Polylactic acid (PLA Tough) used as a construction material. Grid pattern infill of 20% has been set (Fig. 4). The printer was equipped with 0.25AA nozzle. Adhesion plate was added. No secondary support material was provided because none support was needed. The printing lasted 6 hours.

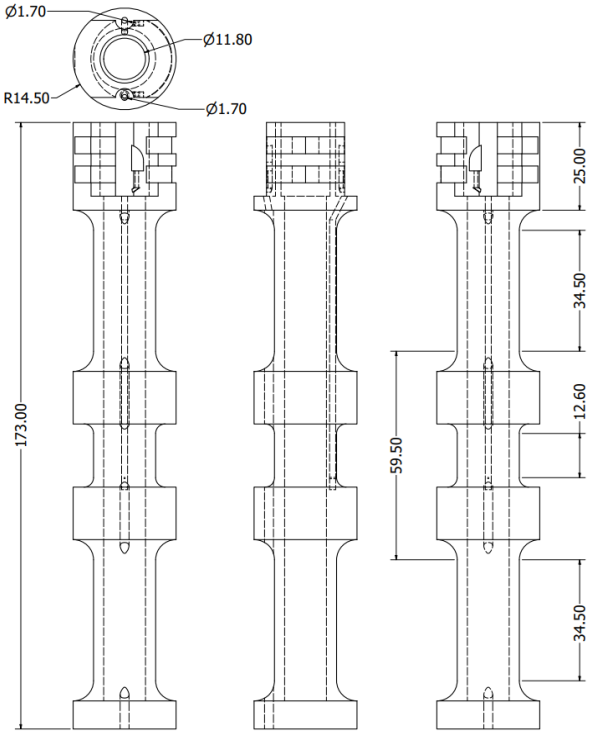


Fig. 3. Technical drawing of carcass designed in Autodesk Inventor Professional 2021. Four projections of the model with marked chosen dimensions on

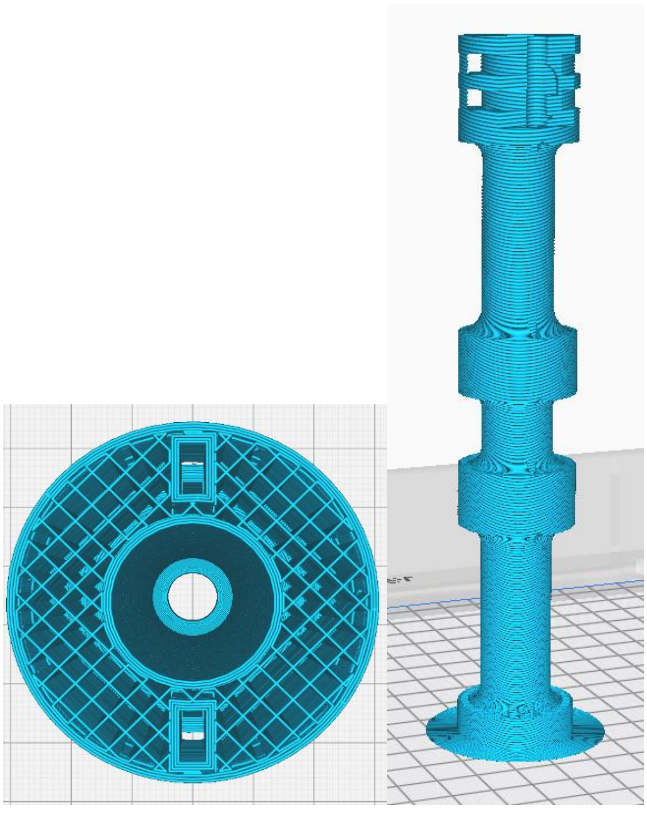


Fig. 4. 3D-view of designed receive coils carcass from Ultimaker Cura – software for setting 3D printing parameters. Infill grid pattern is visible in the cross-section (on the left). Carcass model divided into printable layers (on the right)

1.4. Sample preparation

Samples were prepared using 20 nm iron oxide nanoparticles nanomag-D-spio (Micromod). Dextran iron oxide composite particles were originally dissolved in water. Two samples containing 1 ml of nanoparticle suspension were prepared (10 mm diameter and 17 mm height). The first specimen contained original iron concentration of 2.4 mg/ml as given by the manufacturer. The second one was diluted with demineralised water to 0.5 mg/ml. However, this value should be considered as a rough approximation due to the limited accuracy of the syringe used for drawing the nanoparticle suspension and the diluent. Both samples were measured at room temperature.

1.5. Experimental procedure

The measurement was performed for nine different frequencies, ranging from 2 to 50 kHz using lock-in amplifier.

After assembling the resonant circuit for desired excitation frequency, the signal from the function generator was triggered. Receiving coil system was inserted inside the excitation coil. Empty probe signal was minimized by precise positioning of detection coils in relation to primary excitation field using a manually controlled linear actuator. The procedure had to be repeated for each frequency because the position of minimum differs slightly with frequency. A control measurement of remaining unbalanced signal of empty probe was performed. The signal of reference coil was fed to reference input of lock-in amplifier.

A test tube with nanoparticle solution was then inserted inside the detection coil. The complex magnetization signal of the nanoparticles was measured. The next test tube with the different sample was inserted and a second measurement was performed. Then the resonant circuit had been assembled for next frequency and so the whole procedure was being repeated until all results were collected.

2. Results

2.1. Assessment of constructed coils

To ensure proper functioning of the system the uniformity of magnetic field generated by excitation coil and sensitivity of receive coils was assessed. Uniformity was measured using self-made probing coil connected to an oscilloscope. Results were presented in Fig. 5. The amplitude of the AC magnetic field at 15 kHz was 4.25 mT measured at the centre of excitation coil. Max deviation of uniformity of the magnetic field in the bore of coil was 6.5% at distance of inner 13 cm that corresponds with the length of working volume of receive coils.

Sensitivity was evaluated using thin magnetic sample driven by excitation magnetic field (f = 15 kHz, B = 4.25 mT) and measuring the response signal using receiving coils and the oscilloscope. Results were presented in Fig. 6. The peek value of induced voltage in receive coils differs slightly due to imperfect winding. The receive coils have different winding density – the same number of turns at different lengths (Tab. 2). That results in a remaining unbalanced sinusoidal signal induced in detection coils to less than 3 mV noise level.

2.2. Magnetic susceptibility measurements

The measurements of complex magnetic susceptibility of prepared samples were performed in frequency range of 2.2–50 kHz. The real and imaginary parts of susceptibility were presented in Fig. 7 and Fig. 8 respectively. In-phase magnetic susceptibility value decreases rapidly at low frequencies and significantly slower above 3.3 kHz. On the contrary, out-of-phase magnetic susceptibility grows slowly as frequency increases. Maximum value of out-of-phase component is expected to occur at higher frequencies beyond the range of the experiment where the Brown relaxation can be observed. Nanoparticle concentration in both examined samples had no influence on the shape of magnetic susceptibility function. Iron concentration ratio in examined samples was one to five. However, the ratio of measured magnetic susceptibility values between samples was 3.3 for real part and 3.7 for imaginary component.

Noise in the measured signal was at level of 10 nV but it was not the main factor influencing the measurement uncertainty. When the experiment was repeated, we observed a deterministic error probably caused by changing environmental circumstances (e.g. thermal instability of measuring circuit or high vulnerability to cable displacement). For different time of experiment there was a significant difference in measured values. The discrepancy between values obtained in the different experiment repeated at different times was estimated as an absolute value of the difference between measured values. The value of discrepancy was marked in a form of bars in Fig. 7 and Fig. 8.

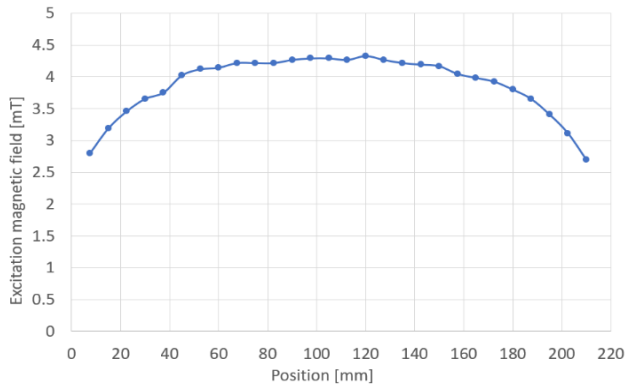


Fig. 5. Spatial distribution of magnetic field amplitude along Z-axis of the excitation coil

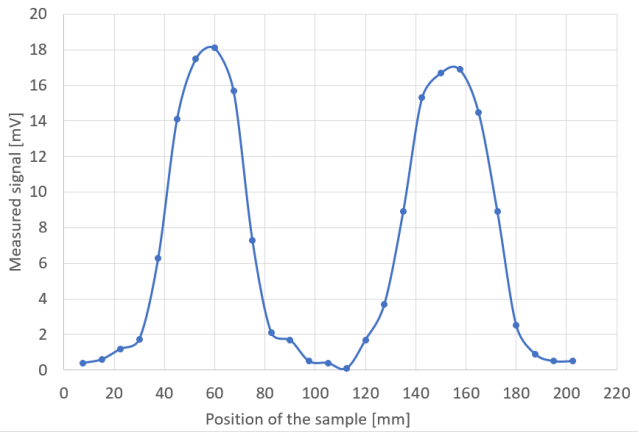


Fig. 6. Distribution of receive coils sensitivity along Z-axis. The value of induced voltage in receive coils in function of small sample position in the coils bore

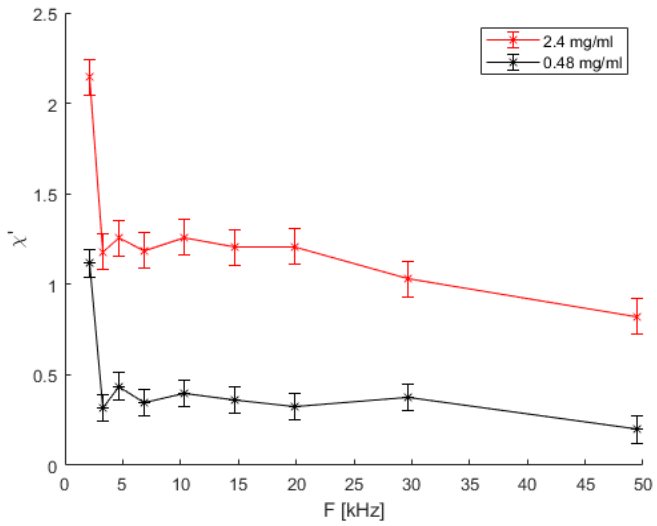


Fig. 7. Real part of magnetic susceptibility in function of excitation frequency. Measurements performed for 2 different concentrations of 20 nm nanomag-D-spio nanoparticles (Micromod Partikeltechnologie GmbH)

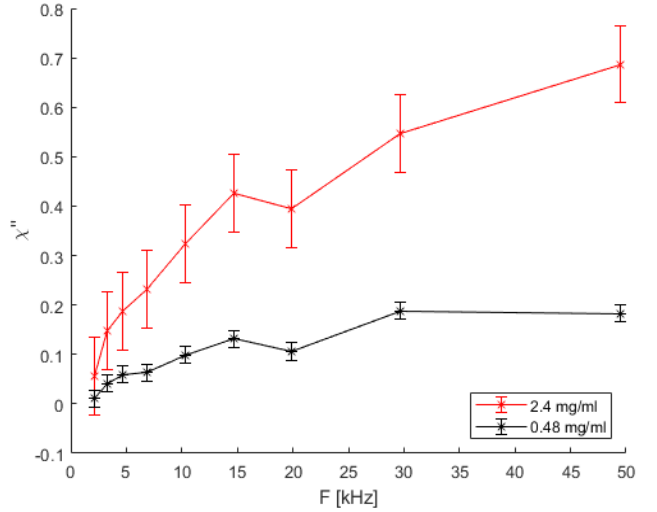


Fig. 8. Imaginary part of magnetic susceptibility in function of excitation frequency. Measurements performed for 2 different concentrations of 20 nm nanomag-D-spio nanoparticles (Micromod Partikeltechnologie GmbH)

3. Conclusions

The laboratory setup for AC impedance spectroscopy of magnetic nanoparticles was designed and built. The receiving coils were fabricated using 3D-printing. Flexibility of 3D-printing allowed us a rapid prototyping of designed carcass. Different material for carcass fabrication were tested. Though polycarbonate (PC) presented the highest maximum operating temperature (121°C) we found that it was too fragile and crumbled during winding. Finally the carcass was made of Tough PLA with maximum working temperature (63°C). Despite the low melting temperature the carcass was able to withstand heat losses generated by current induced due to alternating magnetic field of 1.8 mT.

The uniformity of magnetic field generated by excitation coil in the volume corresponding to the setup of receive coils was sufficient for measurements. Despite the applied configuration of receive coils the complete cancellation of excitation field has not been achieved but was lowered to an acceptable level of unbalance.

Applied symmetrical placement of reference coil in respect to detection coils resulted in the constant value of the phase-angle of the reference signal.

The first measurement using experimental setup was performed in limited frequency range from 2.2 kHz to 50 kHz. Real and imaginary part of magnetic susceptibility in function of frequency obtained in our experiment present a similar characteristic as presented in the literature [1, 7]. Measurements need to be performed in broader spectrum of frequencies for proper estimation of the frequency that maximizes magnetic losses for 20 nm nanomag-D-spio. Other methods such as calorimetric measurements of power losses are planned to be used to confirm the results obtained in the future experiments.

As expected we were able to measure smaller value of magnetic susceptibility for a sample of lower iron concentration. However, the ratio in measured magnetic susceptibility values between both samples was 3.5 in average and did not correspond exactly to iron content differences in prepared samples, which were designed to contain 5 times more iron in the first sample (2.4 mg/dl) compared to the second one (0.48 mg/dl). Other measurements should be made to verify if there is a linear response between measured value of magnetic susceptibility and iron concentration in the examined sample.

Using Lock-In amplifier it is possible to join AC magnetometry (spectrum of different excitation frequencies) with MPS method (spectrum of different harmonics for each excitation frequency). Further research need to be done to investigate combination of MPS-AC magnetometry.

Acknowledgments

This work was supported with grant number: 504/04558/1034/43.050003 awarded by Scientific Council for the Discipline of Biomedical Engineering (RND Inżynieria Biomedyczna) of Warsaw University of Technology.

References

[1] Bogren S. et al.: Classification of Magnetic Nanoparticle Systems–Synthesis, Standardization and Analysis Methods in the NanoMag Project. International Journal of Molecular Sciences 16(9)/2015, 20308–20325 [http://doi.org/10.3390/ijms160920308].
[2] Graeser M. et al.: Analog receive signal processing for magnetic particle imaging. Med. Phys. 40(4)/2013, 042303 [http://doi.org/10.1118/1.4794482].
[3] Harabech M. et al.: The Effect of the Magnetic Nanoparticle's Size Dependence of the Relaxation Time Constant on the Specific Loss Power of Magnetic Nanoparticle Hyperthermia. Journal of Magnetism and Magnetic Materials 426/2017, 206–210 [http://doi.org/10.1016/j.jmmm.2016.11.079].
[4] Hergt R. et al.: Magnetic Particle Hyperthermia: Nanoparticle Magnetism and Materials Development for Cancer Therapy. Journal of Physics Condensed Matter 18(38)/2006, S2919 [http://doi.org/10.1088/0953-8984/18/38/S26].
[5] Kishore K., Akbar S. A.: Evolution of Lock-In Amplifier as Portable Sensor Interface Platform: A Review. IEEE Sensors Journal 20(18)/2020, 10345–10354 [http://doi.org/10.1109/JSEN.2020.2993309].
[6] Ludwig F. et al.: Analysis of AC Susceptibility Spectra for the Characterization of Magnetic Nanoparticles. IEEE Transactions on Magnetics 53(11)/2017, 10–13 [http://doi.org/10.1109/TMAG.2017.2693420].
[7] Mahdavi Z. et al.: Core-Shell Nanoparticles Used in Drug Delivery-Microfluidics: A Review. RSC Advances 10(31)/2020, 18280–18295 [http://doi.org/10.1039/d0ra01032d].
[8] Maity D., Ganeshlenin K.: Superparamagnetic Nanoparticles for Cancer Hyperthermia Treatment. Nanotechnology Characterization Tools for Tissue Engineering and Medical Therapy, Springer Berlin Heidelberg, 2019, 299–332 [http://doi.org/10.1007/978-3-662-59596-1_7].
[9] Reeves D. B., Weaver J. B.: Magnetic Nanoparticle Sensing: Decoupling the Magnetization from the Excitation Field. Journal of Physics D: Applied Physics 47(4)/2013, 45002 [http://doi.org/10.1088/0022-3727/47/4/045002].
[10] Sandler S. E. et al.: Best Practices for Characterization of Magnetic Nanoparticles for Biomedical Applications. Analytical Chemistry 91(22)/2019, 14159–14169 [http://doi.org/10.1021/acs.analchem.9b03518].
[11] Šouc J. et al.: Calibration Free Method for Measurement of the AC Magnetization Loss. Superconductor Science and Technology 18(5)/2005, 592–595 [http://doi.org/10.1088/0953-2048/18/5/003].
[12] Suhaimi N. S. et al.: A Resonant Type AC Magnetometer for Evaluation of Magnetic Nanoparticles. Hassan M. (eds) Intelligent Manufacturing & Mechatronics. Lecture Notes in Mechanical Engineering. Springer, Singapore 2018 [http://doi.org/10.1007/978-981-10-8788-2_9].
[13] Sun Y. et al.: An Improved Method for Estimating Core Size Distributions of Magnetic Nanoparticles via Magnetization Harmonics. Nanomaterials 10(9)/2020, 1–12 [http://doi.org/10.3390/nano10091623].

[14] Valentini M. et al.: Diffusion NMR Spectroscopy for the Characterization of the Size and Interactions of Colloidal Matter: The Case of Vesicles and Nanoparticles. Journal of the American Chemical Society 126(7)/2004, 2142–2147 [http://doi.org/10.1021/ja037247r].
[15] Vallejo-Fernandez G. et al.: Mechanisms of Hyperthermia in Magnetic Nanoparticles. Journal of Physics D: Applied Physics 46(31)/2013 [http://doi.org/10.1088/0022-3727/46/31/312001].
[16] Van De Loosdrecht M. M. et al.: A Novel Characterization Technique for Superparamagnetic Iron Oxide Nanoparticles: The Superparamagnetic Quantifier, Compared with Magnetic Particle Spectroscopy. Review of Scientific Instruments 90(2)/2019 [http://doi.org/10.1063/1.5039150].
[17] Wróblewski P., Smolik W.: Coil design with litze wire for magnetic particle spectrometry. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska 7(1)/2017, 150–153 [http://doi.org/10.5604/01.3001.0010.4605].
[18] Wu K. et al.: Magnetic Particle Spectroscopy: A Short Review of Applications Using Magnetic Nanoparticles. ACS Applied Nano Materials 3(6)/2020, 4972–89 [http://doi.org/10.1021/acsanm.0c00890].
[19] Yang T. Q. et al.: Detection of Magnetic Nanoparticles with Ac Susceptibility Measurement. Physica C: Superconductivity and Its Applications 412–414/2004, 1496–1500 [http://doi.org/10.1016/j.physc.2004.01.146].
[20] Quantum Design, MPMS Application Note 1070-207: Using PPMS Superconducting Magnets at Low Fields 2009.

M.Sc. Eng. Mateusz Midura
e-mail: M.Midura@ire.pw.edu.pl

Mateusz Midura was born in Rzeszów, Poland, in 1994. Received the M. Sc. degree from Warsaw University of Technology, Warsaw, Poland in 2019. He is Ph.D. student in the Nuclear and Medical Electronics Division, IriTM, Electronics and Information Technology Faculty, WUT and the last year student of medicine at Medical University of Warsaw, Faculty of Medicine. His field of interest covers Magnetic Particle Imagining, Magnetic Hyperthermia and position sensitive gamma ray detectors.

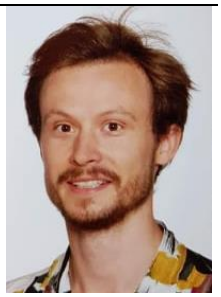
http://orcid.org/0000-0002-2449-0652

M.Sc. Eng. Przemysław Wróblewski
e-mail: P.Wróblewski@ire.pw.edu.pl

Received the M. Sc. degree in biomedical engineering from Warsaw University of Technology, Warsaw, Poland in 2013. Ph.D. student in the Nuclear and Medical Electronics Division, IRiTM, Electronics and Information Technology Faculty, Warsaw University of Technology. His field of interest covers Magnetic Particles Imaging, Magnetic Hyperthermia and Magnetic Particles Spectroscopy, design of scanners and simulation of the magnetic field.

http://orcid.org/0000-0002-6713-9088

M.Sc. Eng. Damian Wanta
e-mail: D.Wanta@ire.pw.edu.pl



Damian Wanta was born in Starogard Gdański, Poland, in 1991. He received the M. Sc. degree in biomedical engineering from Warsaw University of Technology, Warsaw, Poland in 2016. He is Ph.D. student in the Nuclear and Medical Electronics Division, IRiTM, Electronics and Information Technology Faculty, WUT. His current research interests include Imaging of Magnetic Nanoparticles, Electrical Capacitance Tomography and Partial Reconfiguration.

http://orcid.org/0000-0002-1596-6524

Ph.D. Eng. Grzegorz Domański
e-mail: G.Domanski@ire.pw.edu.pl

Received M.Sc. degree and the Ph.D. degree from Warsaw University of Technology, Warsaw, Poland in 1994 and 2001, respectively. Senior lecturer in the Nuclear and Medical Electronics Division, IRiTM, Electronics and Information Technology Faculty, WUT Secretary of Warsaw branch of The Polish Society of Medical Physics (2001 - 2005). Deputy director for radiological protection affairs since 2001. His main research interests are biocybernetics and biomedical engineering.

http://orcid.org/0000-0002-0204-2322

M.Sc. Eng. Mateusz Stosio
e-mail: M.Stosio@ire.pw.edu.pl

Mateusz Stosio was born in Mińsk Mazowiecki, Poland, in 1989. He received the M.Sc. degree in electronics and computer engineering from the Warsaw University of Technology, Warsaw, in 2014. He is currently pursuing Ph.D. degree with the Division of Nuclear and Medical Electronics, Faculty of Electronics and Information Technology, Institute of Radioelectronics and Multimedia Technology. His research interests include impedance measurement and Electrical Capacitance Tomography.

http://orcid.org/0000-0002-7488-1969

Ph.D. Eng. Jacek Kryszyn
e-mail: J.Kryszyn@ire.pw.edu.pl

Jacek Kryszyn was born in Warsaw, Poland, in 1986. He received his M.Sc. degree in electronics and computer engineering and the Ph.D. degree from Warsaw University of Technology, Warsaw, Poland in 2012 and 2018, respectively. He is an assistant professor at the IRiTM, Electronics and Information Technology Faculty, WUT since 2019. His field of interest covers Electrical Capacitance Tomography, especially small capacitance measurement methods.

http://orcid.org/0000-0002-0042-0473

Prof. Waldemar T. Smolik
e-mail: W.Smolik@ire.pw.edu.pl

Waldemar T. Smolik was born in Otwock, Poland, in 1966. He received the M.Sc., the Ph.D. and D.Sc. degrees in electronics engineering from Warsaw University of Technology, Warsaw, Poland in 1991, 1997 and 2014, respectively.
Since 2016, he is a Professor at the Institute of Radioelectronics and Multimedia Technology, Electronics and Information Technology Faculty, WUT. He is the head of the Laboratory of Data Acquisition and Processing Systems at the Division of Nuclear and Medical Electronics.

http://orcid.org/0000-0002-1524-5049

otrzymano/received: 29.01.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2583

MAGNETOELECTRIC COUPLING MEASUREMENT TECHNIQUES IN MULTIFERROIC MATERIALS

Jakub Grotel
Lublin University of Technology, Faculty of Electrical Engineering and Computer Science, Lublin, Poland

Abstract. *Magnetoelectric multiferroics are solid-state materials which exhibit a coupling between ferroelectric and magnetic orders. This phenomenon is known as the magnetoelectric (ME) effect. Multiferroic materials possess a wide range of potential applications in such fields as metrology, electronics, energy harvesting & conversion, and medicine. Multiferroic research is facing two main challenges. Firstly, scientists are continuously trying to obtain a material with sufficiently strong, room-temperature ME coupling that would enable its commercial application. Secondly, the measurement techniques used in multiferroic research are often problematic to implement in a laboratory setting and fail to yield reproducible results. The aim of the present work is to discuss three most commonly used methods in multiferroic studies; the lock-in technique, the Sawyer-Tower (S-T) circuit and dielectric constant measurements. The paper opens with a general description of multiferroics which is followed by mathematical representation of the ME effect. The main body deals with the description of the aforementioned measurement techniques. The article closes with a conclusion and outlook for future research.*

Keywords: multiferroics, magnetoelectric effect, Sawyer-Tower circuit, magnetocapacitance

METODY POMIARU SPRZĘŻENIA MAGNETOELEKTRYCZNEGO W MATERIAŁACH MULTIFERROICZNYCH

Streszczenie. *Magnetoelektryczne multiferroiki należą do grupy materiałów, które wykazują sprzężenie pomiędzy uprządkowaniem ferroelektrycznym a magnetycznym. Zjawisko to nosi nazwę sprzężenia magnetoelektrycznego (ME). Multiferroiki posiadają szereg potencjalnych zastosowań w takich dziedzinach jak metrologia, elektronika, pozyskiwanie i konwersja energii czy medycyna. Badania nad multiferroikami stoją przed dwoma głównymi wyzwaniami. Po pierwsze, naukowcy nieustannie próbują uzyskać materiał o wystarczająco silnym sprzężeniu ME w temperaturze pokojowej, który umożliwiłby jego komercyjne zastosowanie. Po drugie, techniki pomiarowe stosowane w badaniach multiferroików są często problematyczne we wdrażaniu ich w warunkach laboratoryjnych i nie umożliwiają otrzymania powtarzalnych wyników. Celem niniejszej pracy jest omówienie trzech najczęściej stosowanych metod w badaniach nad multiferroikami; techniki lock-in, mostka Sawyera-Towera (S-T) i pomiarów stałej dielektrycznej. Artykuł rozpoczyna się od ogólnego opisu multiferroików, po którym następuje matematyczne przedstawienie efektu ME. W głównej części omówiono wyżej wymienione techniki pomiarowe. Artykuł zamyka podsumowanie i perspektywy przyszłych badań.*

Słowa kluczowe: multiferroiki, efekt magnetoelektryczny, mostek Sawyera-Towera, magnetopojemność

Introduction

Multiferroics are a class of solid-state materials which combine two or more primary ferroic orders simultaneously in the same phase. The primary ferroic orders include ferromagnetism, ferroelectricity and ferroelasticity. Possible interactions between these orders are indicated in Fig. 1 [9, 25].

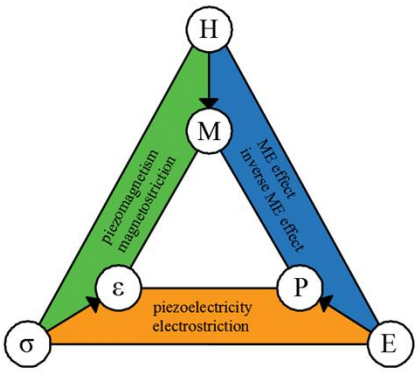


Fig. 1. Interactions between ferroic orders: E – electric field, P – polarization, H – magnetic field, M – magnetization, σ – stress, ε – strain

Magnetoelectric (ME) effect is a coupling between magnetic and ferroelectric orders in a material. It can be expected that exposure of such a multiferroic compound to a magnetic field H will induce an electric polarization P. Conversely, placing the material in an electric field E should result in changes to its magnetization M [9, 14, 25]. If the changes are linear in nature, the strength of this coupling is characterized by the magnetoelectric coupling coefficient α [28].

According to the formal nomenclature, a material that exhibits the ME coupling is called a magnetoelectric multiferroic. Due to immense interest in this form of coupling, the scientific community has reverted to the general name ‘multiferroic’ and nowadays the terms are used interchangeably unless explicitly stated otherwise [9].

The history of multiferroics can be traced back to the late 19th century and the research of Wilhelm Röntgen and Pierre Curie [14, 25]. Further experiments were conducted in the former Soviet Union on boracites in the late 1950s and 1960s by G. Smolenskii, A. Ioffe and Y. Venevtsev. By the 1990s, numerous theories and models had been put forward along with novel experimental techniques which resulted in the revival of this field of study [9, 14, 25]. The term “multiferroic” itself originates from the MEIPIC-2 conference held in 1993 [9]. The next turning point came with the arrival of the 21st century. In 2001, N. Spaldin held the first meeting of the American Physical Society solely dedicated to multiferroic materials. With the discovery of new compounds and mechanisms behind the ME coupling, the number of publications increased exponentially [14, 25]. Nowadays, the field of study is continuing to develop at a dynamic rate with thousands of papers being published every year [29].

For many years, scientists had been baffled by the existence of magnetoelectric multiferroics since the conditions for either ferroic order appear to be mutually exclusive. Ferromagnetism occurs in transition metal compounds, in which cations have their d subshells partially occupied by electrons. The ligands are positioned symmetrically around the metal ion. On the other hand, ferroelectricity favours insulators-systems with empty d subshells and a non-centrosymmetric ligand field surrounding the cation [9, 14, 25]. While the cause of (ferro)magnetism is straightforward and involves exchange interactions between spins and ordering of magnetic moments, the mechanisms of ferroelectricity are more diverse and complicated. These include: ferroelectricity due to lone electron pairs, charge ordering, geometric ferroelectricity and ferroelectricity due to spin-driven mechanisms (e.g., inverse Dzyaloshinskii-Moriya interaction) [4, 9, 14].

Multiferroics can exist as either single-phase compounds or composites and can be classified into two basic groups. In type-I multiferroics, magnetic and ferroelectric orders occur at different temperatures (possibly above room temperature) and their mechanisms are independent of one another. Such materials feature a relatively weak ME coupling. In type-II multiferroics, ferroelectric response is triggered by magnetic phenomena in the material. Although ME coupling tends to be stronger than

in the former case, it is usually achieved at cryogenic temperatures at which non-zero polarization occurs [14]. Despite this drawback, a few novel type II materials have been discovered (e.g., hexaferrites) that yield responses of the applicable order at room temperature [15, 16].

Magnetoelectric multiferroics promise a wide range of potential applications, predominantly in disciplines like physics, materials science and engineering. The materials can be utilized as transducers and magnetic field sensors in metrology, more efficient non-volatile memory (use of electric fields instead of magnetic ones), logic elements, frequency-sensitive capacitors in electronics and spintronics or diodes in photovoltaics. Last but not least, multiferroics can increase the chances of combating cancer via new diagnostics and drug delivery methods in medicine [9, 17, 25, 29].

1. Magnetoelectric coupling coefficient

The ME coupling coefficient can be obtained by examining the free energy F function of a crystal according to the Landau theory of phase transition [1, 8, 10, 22, 24]:

$$F(E,H) = F_0 - P_i^S E_i - M_i^S H_i - \frac{1}{2}\varepsilon_0\varepsilon_{ij}E_iE_j - \frac{1}{2}\mu_0\mu_{ij}H_iH_j - \alpha_{ij}E_iH_j - \frac{\beta_{ijk}}{2}E_iH_jH_k - \frac{\gamma_{ijk}}{2}H_iE_jE_k - \ldots \quad (1)$$

where: F_0 – free energy of a disordered system; P_i^S, M_i^S – spontaneous polarization and magnetization; ε_0 – vacuum permittivity; ε_{ij} – relative permittivity; μ_0 – vacuum permeability; μ_{ij} – relative permeability; α_{ij}, β_{ijk}, γ_{ijk} – magnetoelectric coupling coefficients of successive orders of magnitude [8, 10, 24].

The function F is represented as Maclaurin power series in terms of an external electric field E and magnetic field H. The indices i, j and k correspond to x, y and z axes. Einstein summation convention regarding the same two subscripts in each term is implied [8, 22]. The first term represents free energy of a disordered (high-temperature) phase. The second and third terms quantify energy contributions caused by spontaneous polarization and magnetization, respectively. The fourth and fifth terms account for material response to an electric or magnetic field. The sixth term corresponds to the linear ME coupling. The second-rank tensor α_{ij} stands for the linear ME coupling coefficient, whereas higher-order (quadratic) ME effects are denoted by third-rank tensors β_{ijk} and γ_{ijk}. These tensor coefficients are measured in [sm^{-1}] in SI units [8, 10, 22]. Unless it is specified, the ME coupling under investigation should be understood as a linear effect.

By differentiating equation (1) with respect to E_i or H_i and then setting E_j and H_j respectively as 0, the following ME responses are obtained [8, 10, 22]:

$$P_i(H) = P_i^S + \alpha_{ij}H_j + \frac{\beta_{ijk}}{2}H_jH_k + \ldots \quad (2a)$$

$$M_i(E) = M_i^S + \alpha_{ji}E_j + \frac{\gamma_{ijk}}{2}E_jE_k + \ldots \quad (2b)$$

An important difference should be noted between these equations, namely the change $\alpha_{ij} \rightarrow \alpha_{ji}$. The latter coefficient quantifies the (linear) inverse ME coupling.

Finally, a clear representation of the ME coupling coefficient is obtained with further differentiation:

$$\alpha_{ij}^H = \frac{\partial P_i}{\partial H_j} \quad (3a)$$

$$\alpha_{ji}^E = \frac{\partial M_i}{\partial E_j} \quad (3b)$$

The superscripts inform as to the source of the ME coupling. The second rank tensor α_{ij} takes the form of a 3×3 matrix:

$$\alpha_{ij}^H = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix} \equiv \alpha \quad (4)$$

In most cases, only one out of nine components is non-zero due to crystal symmetry of the sample and external fields geometry considerations [28]. In polycrystalline multiferroics, the ME coupling coefficient is a scalar quantity [28]. The extent of the ME response is then limited by the condition:

$$\alpha^2 \leq \varepsilon_0\mu_0\varepsilon_r\mu_r \quad (5)$$

where: ε_r – relative permittivity, μ_r – relative permeability [8, 28].

An alternative, thermodynamic approach to multiferroic single-phase and composite systems can be found in [29].

2. Measurement techniques

There are two challenges that dominate the multiferroic research. Firstly, a material with sufficiently strong, room-temperature ME coupling for large-scale commercial use has not been identified yet. Secondly, the measurement techniques used in multiferroic research are often problematic to implement in a laboratory setting and, despite being thoroughly studied, fail to yield reproducible results. The known experimental methods are classified as direct or indirect. Direct methods assign a numerical value to the ME coupling coefficient, whereas indirect methods just inform about the presence of the phenomenon in the measured sample. Furthermore, indirect methods are usually insufficient to determine multiferroicity on their own and require additional studies [3, 28]. The following subsections describe three most commonly used methods in multiferroic studies; the lock-in technique, the Sawyer-Tower (S-T) circuit and dielectric constant measurements.

2.1. Lock-in technique

The lock-in technique, also known as the dynamic method, takes its name from the key component of the experimental set-up – a lock-in amplifier. In this method, a multiferroic sample (mono- or polycrystalline) is exposed to a sinusoidal, alternating current (AC) magnetic field with an additional direct current (DC) bias. The superimposed fields lead to electric polarization of the material via the magnetoelectric effect. The resultant voltage V across the sample is measured by the frequency-attuned lock-in amplifier. The ME coupling coefficient is then calculated based on the gathered data [19, 21, 22, 28].

The experimental set-up consists of the following components:

- a pair of Helmholtz coils,
- electromagnet with a power source,
- Hall probe,
- function generator,
- lock-in amplifier,
- PC with LabView software,
- multimeter,
- sample [7, 11].

The way the different components are connected is presented in Fig. 2.

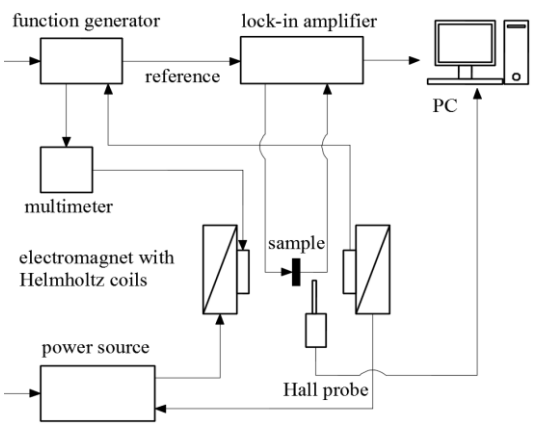


Fig. 2. The lock-in technique experimental set-up

A pellet-shaped sample is placed inside a pair of Helmholtz coils, between the poles of a DC electromagnet. The sample is oriented perpendicular to the magnetic field lines and its position ensures excitation filed uniformity. A Hall probe placed next to the sample measures the intensity of the DC bias field. The field typically takes values from a range of ±15 kOe (1 Oe = 79.577 Am^{-1} in SI units) [7, 11, 18, 28]. The Helmholtz coils connected to a function generator generate an AC field at frequency *f* (usually up to 10 kHz) and intensity of about 20 Oe. The exact field intensity can be calculated with an appropriate formula by measuring the driving current in the circuit. The sample is connected via electrodes to the lock-in amplifier operating in the voltage differential mode.

The lock-in amplifier is an instrument used for tracking and amplification of weak signals in the presence of background noise. Its working principle is simple. The device utilizes a phase-sensitive detection technique to pick out and follow signals with a given frequency and phase shift. This is achieved by comparing a noisy input signal with a reference signal (e.g., from a function generator) continuously in real-time. The noisy voltage signal is first amplified and passed through a bandpass filter. The lock-in amplifier tracks the signal parameters and makes adjustments to the reference signal accordingly. Both signals are multiplied by a special rectifier- the phase-sensitive detector. If the frequencies and phases match, a DC signal proportional to the original is obtained. The signal is then passed through a low pass filter and amplified. The values can be read directly from the device or processed with a computer interfaced with the amplifier (data acquisition card required) [30].

Finally, the voltage ME coupling coefficient is calculated from the collected data:

$$\alpha_V^H = \frac{1}{t}\left(\frac{\partial V}{\partial H}\right) = \frac{V_\omega}{t h_0} = \frac{\alpha + 2\beta H + 3\gamma H^2 + \ldots}{t} \quad (6)$$

where: t – sample thickness, V_ω – first harmonic of the magnetically induced voltage, h_0 – AC magnetic field amplitude, H – DC magnetic field amplitude.

The unit of the voltage coefficient is $Vcm^{-1}Oe^{-1}$ [28]. Equation (6) leads to an important observation. If a DC bias is used in the experiment, the calculated coefficient is not truly linear. In fact, it is called a pseudo-linear coefficient which value depends on the contribution from higher-order ME coefficients [22]. In reality, an AC excitation field alone is usually not enough to conduct the measurements [28].

Despite the amplifier ability to deal with noisy signals, noise can have a meaningful impact on the outcome of the experiment. It is always important to identify the sources of noise and eliminate or reduce their influence. Noise can be intrinsic (inherent to all physical processes) or external (due to environmental factors) [30].

Types of noise:

1) Intrinsic noise sources:
- Johnson noise – an electronic noise due to thermal agitation of charge carries in a conducting material [30],
- shot noise (granular noise) – an electronic noise caused by the discrete nature and statistical independence of individual charge carriers [30, 31],
- 1/f (flicker noise) – an electronic noise which is presumably linked to fluctuating domains and defects in a material structure [30, 32],

2) Extrinsic noise sources:
- capacitive coupling – parasitic capacitance between electrical conductors. Two parallel conductors at different potentials and in close proximity act as capacitors plates [30, 33],
- inductive coupling – a noise voltage generated by an AC in one conductor which induces an EMF (electromotive force) in a nearby second conductor via Faraday's law of induction [30, 33],
- ground loops – multiple ground points in an electrical circuit can create a current loop that picks up noise [30],
- microphonics – mechanical vibrations cause a noise current in an electrical circuit [30],
- thermocouple effects – a junction of two dissimilar metals generates a thermoelectric EMF if the metals are at different temperatures [30].

2.2. Sawyer-Tower circuit

The Sawyer-Tower (S-T) circuit is an indirect method of multiferroic investigation. Originally, the circuit is a simple way to characterize dielectric materials – observe and measure their electrical hysteresis under variable electric field intensity [5, 6]. The electric polarization in multiferroics is non-linear, therefore the confirmation of the ME effect can be readily made with minor modifications to the original circuit [28]. The method should be treated as a tool for preliminary ferroelectric and multiferroic studies.

The basic Sawyer-Tower circuit consists of the following components:
- function generator,
- high-voltage (HV) amplifier,
- resistive voltage divider,
- capacitive voltage divider,
- compensation circuit,
- oscilloscope,
- data acquisition (DAQ) card,
- PC,
- sample [5, 26].

The following schematic (Fig. 3) shows the basic S-T circuit experimental set-up.

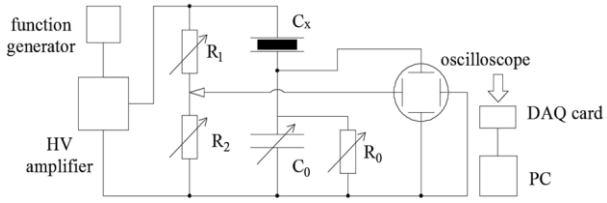


Fig. 3. The Sawyer-Tower circuit experimental set-up

In the basic design, a signal generator with a high-voltage amplifier supplies the S-T circuit with a sinusoidal signal. The circuit consists of two voltage dividers; resistive (R_1 and R_2) and capacitive (a measured sample C_x and a decade capacitor C_0). The output voltage of the first divider registered by the horizontal plates of the oscilloscope. The voltage across the capacitor C_0 is measured by the vertical plates. The oscilloscope works in XY mode and the perpendicular signals result in a two-dimensional, real-time function $P(E)$ being displayed on the oscilloscope screen. The circuit might be additionally equipped with a compensation to counter the effects of dielectric losses (frequency- and capacitance-dependent dissipation of energy of a material [5]. The compensation follows the formula:

$$R_0 C_0 = R_x C_x \quad (7)$$

where R_x – resistance of the sample.

The data can be processed further by means of suitable software. A connection between the experimental set-up and a computer can be established with a DAQ card [5, 26].

Each dielectric material has a parameter known as dielectric strength. It is a field intensity limit, exceeding which leads to dielectric breakdown and damage to the sample. In multiferroic research, field intensity up to 10 $kVcm^{-1}$ is usually required [1]. The field can be calculated by dividing the voltage across the sample by its thickness. The voltage U across the HV amplifier terminals is measured with a voltmeter and is equal to:

$$U = U_x + U_0 = U_x + \frac{C_x}{C_0}U_x = U_x\left(1 + \frac{C_x}{C_0}\right) \quad (8)$$

where U_x – voltage across the sample, U_0 – voltage across the capacitor C_0.

Since $C_x \ll C_0$, it follows that $U \approx U_x$.

As far as multiferroic studies with the S-T circuit are concerned, the experimental process relies on the mono- or polycrystalline sample being additionally placed between the poles of an electromagnet. Polarization information regarding the ME effect is acquired by changing the intensity of a magnetic field. Since the method is based on visual observation rather than obtaining numerical values, it is important to distinguish between valid results and artifacts. In some cases, a hysteresis might appear to be deformed or be in fact a Lissajous figure (the sample is not multiferroic). One such negative contribution to the experiment is the aforementioned dielectric loss [6].

2.3. Dielectric constant measurements

Dielectric constant (relative permittivity) measurements constitute an alternative method of detecting the magnetoelectric (ME) effect. This indirect method is based on the magnetodielectric effect (MDE) which describes a relationship between magnetic induction B and relative permittivity ε of a material [2, 3, 20, 27]. Since a multiferroic compound relies on a magnetic field to influence its electric polarization, observable changes to its dielectric constant are to be expected.

Relative permittivity is defined as a complex number containing a real and an imaginary part:

$$\varepsilon(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega) \tag{9}$$

where: ε', ε'' – real and imaginary part of relative permittivity, ω – electric field frequency [2, 27].

The Argand diagram showing components of relative permittivity on a complex plane is presented in Fig. 4.

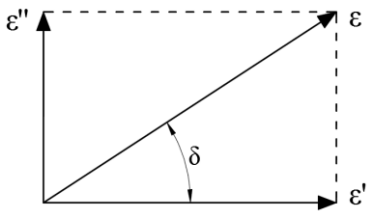


Fig. 4. Complex relative permittivity – the Argand diagram

The real component of complex relative permittivity is in phase with the excitation signal and corresponds to the energy storage capability of the material, whereas the imaginary component is out of phase and is proportional to the loss tangent tanδ (dielectric losses) [1].

Complex relative permittivity is measured with an LCR meter or impedance analyser. The experimental set-up consists of the following components:

- LCR meter,
- sample holder (cell),
- sample [2, 12].

Once the sample in the form of a thin plate is connected to the LCR meter via electrodes, the basic dielectric measurements can commence. A dedicated sample holder with shielding may or may not be used to cut off external noise sources from the experiment.

A modern meter works on a principle of radio frequency current-voltage (RF I-V) measurements [34]. This allows for greater precision over a wider frequency spectrum. The meter generates a sinusoidal voltage in its internal oscillator which is then applied across the sample. The potential difference between the two electrodes, the electric current passing through them and the phase difference δ between these quantitates is measured. The LCR meter derives the value of impedance from acquired data which can also be presented in terms of inductance L, capacitance C and resistance R with the use of a phase-sensitive detector [34, 35]. For high-accuracy impedance measurements, Kelvin (4-wire) method is often utilized to eliminate errors due to wire resistance [35].

There are numerous methods of measuring relative permittivity with an LCR meter. Depending on the available equipment, effective testing standards and required precision, each solution has respective guidelines and formulas to follow. Their description is beyond the scope of this article. The contacting and non-contacting electrode (air-gap) methods are the two simplest ways of obtaining the desired quantity. In essence, they both involve two capacitance and tangent loss measurements; one with the dielectric inserted between the meter electrodes and one without the dielectric. Such a procedure mirrors the physical interpretation of relative permittivity. From a theoretical standpoint, the real part of complex relative permittivity is equal to:

$$\varepsilon' = \frac{C_x}{C_0} \tag{10}$$

where C_x – capacitance with an inserted dielectric, C_0 – capacitance of air or vacuum [1, 23].

Based on Fig. 4, the imaginary part is calculated from the following trigonometric identity:

$$\varepsilon'' = \varepsilon' \tan\delta \tag{11}$$

In order to test a multiferroic materials for the presence of the MDE effect, a sample is additionally placed between the evenly-spaced poles of an electromagnet. The measurements are carried out with the static magnetic field turned on and off. The data is required to calculate the following parameters:

- magnetocapacitance (MC):

$$MC = \frac{\varepsilon'(B) - \varepsilon'(0)}{\varepsilon'(0)} \tag{12}$$

where: $\varepsilon'(B)$, $\varepsilon'(0)$ – relative permittivity measured under an AC electric field with a magnetic field turned on and off [3, 27].

- magnetolosses (ML):

$$ML = \frac{\tan\delta(B) - \tan\delta(0)}{\tan\delta(0)} \tag{13}$$

where: tanδ(B), tanδ(0) – loss tangent measured under an AC electric field with a magnetic field turned on and off [3, 27].

- magnetoresistance (MR):

$$MR = \frac{\Delta\rho}{\rho} = \frac{\rho(B) - \rho(0)}{\rho(0)} = \frac{R(B) - R(0)}{R(0)} \tag{14}$$

where: ρ – resistivity, R – electrical resistance [2, 3, 13].

Average experimental conditions are as follows: sample thickness: 1 mm; frequency range: $10^2 \div 10^6$ Hz; magnetic field strength: ±10 kOe and electric field strength: 1 Vmm^{-1} [2, 12, 27].

The largest single drawback of relative permittivity measurements is that they are not completely reliable when it comes to confirming the occurrence of the ME effect. The MDE effect can take place in materials which are not multiferroic in nature. Magnetocapacitance may have a non-zero value due to a number of intrinsic and extrinsic contributions, such as: sample inhomogeneity, change of polarization of oxygen octahedra owing to interaction between a magnetic field and Fe ions, antiferromagnetic spin fluctuations, etc. [2, 3, 27]. The most common factors causing ME-unrelated magnetocapacitance are magnetoresistance and the creation of interfaces (e.g., grain boundaries) in singe crystals, polycrystalline and composite samples. The latter phenomenon is described and modelled by the Maxwell-Wagner effect [2, 3].

3. Conclusion

Multiferroics are a group of electrical insulators which feature interesting magnetic and electrical properties. The magnetoelectric effect can be used in a variety of disciplines ranging from energy harvesting to medical applications. The perspectives on the future of multiferroic research appear to be brighter than ever. There is a great deal of pressure to discover and commercialize new materials. One can easily observe global trends regarding clean energy, advancements in technologies like quantum computing

or rising rates of cancer or other debilitating conditions that demand new diagnostics and treatment methods.

The lock-in technique, Sawyer-Tower circuit and dielectric measurements are the simplest and most affordable measurement techniques in multiferroic research. Problems with accuracy and result reproducibility show that more emphasis should be put on novel experimental techniques by the scientific community. New multiferroics exhibiting previously unknown phenomena encourage constant adaptation of the existing methods and innovation.

References

[1] Bain A. K., Chand P.: Ferroelectrics: Principles and Applications. Wiley, 2017
[2] Bonaedy T, Koo Y. S., Sung K. D., Jung J. H.: Resistive magnetodielectric property of polycrystalline γ-Fe_2O_3. Applied Physics Letters 91(13)/2007, 132901 [http://doi.org/10.1063/1.2790474].
[3] Catalan G.: Magnetocapacitance without magnetoelectric coupling. Applied Physics Letters 88(10)/2006, 102902 [http://doi.org/10.1063/1.2177543].
[4] Cheong S.-W., Mostovoy M.: Multiferroics: a magnetic twist for ferroelectricity. Nature Materials 6(1)/2007, 13–20 [http://doi.org/10.1038/nmat1804].
[5] Das C., Shahee A., Lalla N., Shripathi T.: A simple and low cost Sawyer-Tower ferro-electric loop tracer with variable frequency and compensation circuit. Proceedings of the 54th DAE Solid State Physics Symposium, 2009, 439.
[6] Dawber M., Rabe K. M, Scott J. F.: Physics of thin-film ferroelectric oxides. Reviews of Modern Physics 77(4)/2005, 1083 [http://doi.org/10.1103/RevModPhys.77.1083].
[7] Duong G. V., Groessinger R., Schoenhart M., Bueno-Basques D.: The lock-in technique for studying magnetoelectric effect. Journal of Magnetism and Magnetic Materials 316(2)/2007, 390-393 [http://doi.org/10.1016/j.jmmm.2007.03.185].
[8] Eerenstein W., Mathur N. D., Scott J. F.: Multiferroic and magnetoelectric materials. Nature (7104)/2006, 759–765 [http://doi.org/10.1038/nature05023].
[9] Fiebig M., Lottermoser T., Meier D., Trassin M.: The evolution of multiferroics. Nature Reviews Materials 1/2016, 16046 [http://doi.org/10.1038/natrevmats.2016.46].
[10] Fiebig M.: Revival of the magnetoelectric effect. Journal of Physics D: Applied Physics 38(8)/2005, R123 [http://doi.org/10.1088/0022-3727/38/8/R01].
[11] Fuentes-Cobas L. E., Matutes-Aquino J. A., Fuentes-Montero M. E.: Handbook of Magnetic Materials, Chapter 3- Magnetoelectricity. Elsevier, 2011
[12] Guobin C., Hui Y., Xiaoming Z., Jun L., Jun T.: Clarification of the Magnetocapacitance Mechanism for Fe_3O_4-PDMS Nanocomposites. Journal of Nanomaterials 2015/2015, 982174 [http://doi.org/10.1155/2015/982174].
[13] Hishiyama Y., Kaburagi Y., Inagaki M.: Materials Science and Engineering of Carbon: Characterization. Elsevier, 2016
[14] Khomskii D.: Classifying Multiferroics: Mechanisms and Effects. Physics 2/2009 [http://doi.org/10.1103/Physics.2.20].
[15] Kimura T.: Magnetoelectric Hexaferrites. Annual Review of Condensed Matter Physics 3(1)/2012, 93–110 [http://doi.org/10.1146/annurev-conmatphys-020911-125101].
[16] Kitagawa Y., Hiraoka Y., Honda T., Ishikura T., Nakamura H., Kimura T.: Low-field magnetoelectric effect at room temperature. Nature Materials 9(10)/2010, 797–802 [http://doi.org/10.1038/nmat2826].
[17] Kreisel J., Kenzelmann M.: Multiferroics – the challenge of coupling magnetism and ferroelectricity. Europhysics News 40(5)/2009, 17–20 [http://doi.org/10.1051/epn/2009702].
[18] Kuila S., Tiwary S., Sahoo M. R., Barik A., Vishwakarma P. N.: Measurement of temperature dependent magnetoelectricity in $BiFe_{(1-x)}Co_xO_3$; x = 0, 0.01, 0.02. Journal of Alloys and Compounds 709/2017, 158–164 [http://doi.org/10.1016/j.jallcom.2017.03.118].
[19] Mahesh Kumar M., Srinivas A., Suryanarayana S. V., Kumar G. S., Bhimasankaram T.: An experimental setup for dynamic measurement of magnetoelectric effect. Bulletin of Materials Science 21(3)/1998, 251–255 [http://doi.org/10.1007/BF02744978].
[20] Parish M. M.: Magnetocapacitance without magnetism. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 372(2009)/2014 [http://doi.org/10.1098/rsta.2012.0452].
[21] Rivera J. P.: A short review of the magnetoelectric effect and related experimental techniques on single phase (multi-) ferroics. The European Physical Journal B 71/2009, 299 [http://doi.org/10.1140/epjb/e2009-00336-7].
[22] Rivera J. P.: On definitions, units, measurements, tensor forms of the linear magnetoelectric effect and on a new dynamic method applied to Cr-Cl boracite. Ferroelectrics 161(1)/1994, 165–180 [http://doi.org/10.1080/00150199408213365].
[23] Serway R., Jewett Jr. J. W.: Physics for Scientists and Engineers. Cengage Learning, 2014.
[24] Siratori K., Kohn K., Kita E.: Magnetoelectric Effect in Magnetic Materials. Acta Physica Polonica A 81/1992, 431–466 [http://doi.org/10.12693/APhysPolA.81.431].
[25] Spaldin N. A., Cheong S.-W., Ramesh R.: Multiferroics: Past, present, and future. Physics Today 63(10)/2010, 38–43 [http://doi.org/10.1063/1.3502547].
[26] Stewart M., Cain M., Hall D.: Ferroelectric Hysteresis Measurement and Analysis. NPL Report 152, 1999.
[27] Turik A. V., Pavlenko A. V.: Magnetodielectric Effect and Magnetoelectricity in Multiferroics and Heterogeneous Systems: Modeling and Experiment. Ferroelectrics 444(1)/2013, 53–59 [http://doi.org/10.1080/00150193.2013.786308].
[28] Vopson M. M., Fetisov Y. K., Caruntu G., Srinivasan G.: Measurement Techniques of the Magneto-Electric Coupling in Multiferroics. Materials 10(8)/2017 [http://doi.org/10.3390/ma10080963].
[29] Vopson M. M.: Fundamentals of Multiferroic Materials and Their Possible Applications. Critical Reviews in Solid State and Materials Sciences 40(4)/2015, 223–250 [http://doi.org/10.1080/10408436.2014.992584].
[30] Stanford Research Systems, Inc., Model SR830 DSP Lock-In Amplifier. User Manual, 2011.
[31] http://physicsopenlab.org/2016/10/10/shot-noise-and-electron-charge/ (available: 02.2021).
[32] http://www.scholarpedia.org/article/1/f_noise (available: 02.2021).
[33] https://techweb.rohm.com/knowledge/emc/s-emc/01-s-emc/6943 (available: 02.2021).
[34] https://www.electronics-notes.com/articles/test-methods/lcr-meter-bridge/primer-basics.php (available: 02.2021)
[35] https://www.voltech.com/Products/DC1000A/Howitworks.aspx (available: 02.2021).

M.Sc. Eng. Jakub Grotel,
e-mail: j.grotel@pollub.pl

Jakub Grotel is an electrical engineering graduate and a Ph.D. student at the Lublin University of Technology. Since 2018, the author has been conducting research in the Department of Electronics and Information Technology under the supervision of Prof. Elżbieta Jartych and Dr. Tomasz Pikula. His interests include materials science, renewable energy engineering and physics. He is the co-author of articles and teaching materials for students.

http://orcid.org/0000-0001-8428-2292

otrzymano/received: 01.03.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2576

METHODS FOR DETECTING FIRES IN ECOSYSTEMS USING LOW-RESOLUTION SPACE IMAGES

Valerii Shvaiko, Olena Bandurka, Vadym Shpuryk, Yevhen Havrylko
The National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Heat Power Engineer Department, Automation of Projection of Power Processes and Systems, Kyiv, Ukraine

Abstract. *The paper presents the methods for fire identification using low-resolution space images obtained from Terra Modis and NOAA satellites. There are lots of algorithms to identify potentially "fire pixels" (PF). They are based on the assessment of temperature in spectral ranges from 3.5–4 to 10.5–11.5 microns. One of the problematic aspects in the Fire Detection Method using low-resolution space images is "Cloud and Water Masking". To identify "fire pixels", it is important to exclude from the analysis fragments of images that are covered with clouds and occupied by water objects. Identification of pixels in which one or more fires are actively burning at the time of passing over the Earth is the basis of the algorithm for detecting potentially "fire pixels". The algorithm requires a significant increase in radiation in the range of 4 micrometers, as well as on the observed radiation in the range of 11 micrometers. The algorithm investigates each pixel in a scene that is assigned one of the following classes as a result: lack of data, cloud, water, potentially fire or uncertain. The pixels that lack actual data are immediately classified as "missing data (NULL)" and excluded from further consideration. Cloud and water pixels, defined by the cloud masking technique and water objects, belong to cloud and water classes, respectively. The fire detection algorithm investigates only those pixels of the Earth's surface that are classified as potentially fire or uncertain. The method was implemented using the Visual Programming Tool PowerBuilder in the data processing system of Erdas Imaging. As a result of the use of the identification method, fires in the Chornobyl exclusion zone, steppe fires and fires at gas wells were detected. Using the method of satellite fire identification is essential for the prompt detection of fires for remote forests or steppes that are poorly controlled by ground monitoring methods.*

Keywords: environmental security, ecosystem fires in Ukraine, remote sensing, GIS

METODY WYKRYWANIA POŻARÓW W EKOSYSTEMACH PRZY UŻYCIU ZDJĘĆ SATELITARNYCH O NISKIEJ ROZDZIELCZOŚCI

Streszczenie. *W artykule przedstawiono metody identyfikacji pożarów przy wykorzystaniu niskorozdzielczych zdjęć satelitarnych uzyskanych z satelitów Terra Modis i NOAA. Istnieje wiele algorytmów służących do identyfikacji potencjalnych "pikseli pożaru". Opierają się one na ocenie temperatury w zakresach spektralnych od 3,5–4 do 10,5–11,5 mikronów. Jednym z problematycznych aspektów metody detekcji pożarów z wykorzystaniem zdjęć satelitarnych o niskiej rozdzielczości jest "maskowanie przez chmury i wodę". Aby zidentyfikować „piksele pożaru" należy wykluczyć z analizy fragmenty zdjęć pokryte chmurami oraz zajęte przez obiekty wodne. Podstawą algorytmu do wykrywania potencjalnych „pikseli pożaru" jest identyfikacja pikseli, w których, w momencie przelotu nad Ziemią, aktywny jest jeden lub większa liczba pożarów. Do prawidłowej pracy, algorytm wymaga znacznego wzrostu poziomu promieniowania w zakresie 4 oraz 11 mikrometrów. Algorytm analizuje każdy piksel zdjęcia i w rezultacie przypisuje mu jedną z następujących klas: brak danych, chmura, woda, potencjalny pożar lub niepewny. Piksele, którym nie są przypisane rzeczywiste dane, są natychmiast klasyfikowane jako "brakujące dane (NULL)" i wyłączane z dalszej analizy. Piksele należące do chmur i wody, określone dzięki technice maskowania chmur i obiektów wodnych, zaliczane są odpowiednio do klas chmur i wody. Algorytm detekcji pożaru bada tylko te piksele powierzchni Ziemi, które zostały zaklasyfikowane jako potencjalny pożar lub niepewne. Metoda została zaimplementowana przy użyciu Visual Programming Tool PowerBuilder w systemie przetwarzania danych firmy Erdas Imaging. W wyniku zastosowania metody, wykryto pożary w czarnobylskiej strefie zamkniętej, pożary stepów oraz pożary przy odwiertach gazowych. Zastosowanie metody satelitarnej identyfikacji pożarów jest niezbędne do szybkiego wykrywania pożarów w odległych lasach lub stepach, które są słabo kontrolowane przez naziemne metody monitoringu.*

Słowa kluczowe: bezpieczeństwo środowiska, pożary ekosystemów na Ukrainie, zdalne wykrywanie, GIS

Introduction

One of the most dangerous natural phenomena is fires in ecosystems (FE). The most difficult and hazardous are the fires in forests, forest-steppes, steppes, leading to catastrophic consequences to the environment. The increase in the number and scale of such fires, linked to the global warming processes, was not spared Ukraine. Ukraine and Poland have common borders and so we have the same problem.

Ukraine has some experience in assessing such fires.

One way of resolving the technical task on collecting and assessing fire information is Earth remote sensing. This has led to active application of (ERS) methods to obtain fire information, solve above-mentioned tasks in Ukraine. The constant evolution of technical means, ways, high frequency of information, availability and accessibility of (ERS) materials enable us to continue research in this direction and solve interesting scientific problems.

1. Formulation of the problem

Space Techniques for Monitoring and Fires in ecosystems help identify them at an early stage and ensure rapid decision – making that facilitates the monitoring and impact assessment. The use of space-based data for fire monitoring provides objective and independent information in a timely and cost-effective manner in order to make quick decisions to subdue the elements. The use of space-based imagery not only allows for the detection of fires in some territories but to perform their primary classification according to the scale of combustion, and the use of multispectral data in bands combinations – is to monitor smoke fields of contamination, assess the transport of combustion products.

There are two main sources of fire activity, which are characterized by specific approaches: combustion sources and fire-damaged territories. But the approach to the fire detection is characterized by certain restrictions and can`t be used for spatial fire coverage evaluation, especially since the satellite is often not located at the point necessary for detection at the moment of active combustion, or it is closed from it due to cloudiness. Therefore, there is a need to obtain data on the territories damaged by fire.

Late detection of forest fires have serious consequences for the natural balance of ecosystems. So monitoring of potentially hazardous areas makes it easier to localize and fight fire in an early stage. Remote monitoring of forest fires is rather promising at the present stage. The data on conditions on Earth being obtained from satellites on a regular basis is widely used for operational monitoring of wildfire. At the same time the possibilities of modern technologies make it possible to combine a large amount of other related information.

2. Analysis of recent research and publications

The detailed analysis of recent study and results indicate that the task of defining areas damaged by (FE) through the use of remote-sensing has many alternative solutions.

There are a lot of online services of fire monitoring through the use of remote-sensing data, such as FIRMS [1].

The following satellite systems are being used for monitoring areas damaged by fires:

- Terra satellites with spectra radiometer
- LANDSAT TM/ETM+ satellites.
- NOAA/AVHRR satellites.

Every satellite system should be used for different purposes. LANDSAT is successfully used for detection and control forest fires, for detection of effect of wind storm on forest planting. NOAA and TERRA are important for detection of dangerous organism concentration [2].

Objective and timely information on fire impact is necessary for a wide range of forestry applications, including the conservation and protection of forests, management of forests, and reforestation updating of forest resources data [5]. Data on forest fires cover is a consistent part of information support assessment of direct and secondary pyrotechnic emissions of carbon dioxide and other greenhouse gases to the atmosphere as part of commitments of countries to the Kyoto Protocol, as well as their scientific research on the problems of global climate change and biosphere [4].

Satellite imagery has been used with increased frequency recently that allows to significantly improve the completeness of the detection and the accuracy of defining the area of tree formations, that dry up, which can be a fire risk, especially in summer, as well to more effectively organize ground-based assessment of quantitative characteristics of their conditions [2, 10].

Possibility of using satellite data for evaluating the extent of damage caused by forest fires. The research findings in particular showed correlation between the damage level values and the display coefficient value in medium infrared band [10, 12], normalized difference vegetation index and a number of other indices derived from space images. The usefulness of combining different pre- and -post satellite data was noted. Most studies up to now have been local, and were carried out in small areas, often limited to individual fires.

Automatic technologies of satellite monitoring which have been developed up for now ensure a continuous detection and assessment of forest fire damage trough integrated use of MODIS and Landsat –TN/ETM, +data [2].

The use of satellite systems became possible thanks to the use of devices that capture light from different spectra including infrared ones.

Aerospace monitoring provide up-to-date information which is particularly important for conducting a situational analysis for the purpose of making the best decisions. These data are the basis for the development of topographic and thematic maps, actually the primary source of all up-to-date cartographic information.

Moreover, modern technology for aerospace monitoring and remote sensing and computer processing held a great deal of advantage over traditional paper charts-on the content and diversity of visualization methods. According to experts` evaluation in the very near term (ERS) Earth Remote Sensing is going to become the main information source for (GIS) while traditional maps will be used only in the initial phase as a source of static information (relief, hydrography, main roads, administrative division). You can also add that practically all computer-based geographic analysis is performed with data in raster form which is characteristic of Earth Remote Sensing. In this study NOAA-AVHRR space imagery were used and are requested by national fire services of many countries (the USA, Canada, Brazil) and have long been used to identify fires and assess their consequences. Despite the fact that the resolution image capability is relatively small – 1100 m, monitoring of heat changes from a pixel makes it possible to identify even forest fires within a small area in hectars, and high-temperature point fire sources. Use the experience of the British Institute of Natural Resources and academic centres of Poland, Chile, Nicaragua, Namibia made it possible to develop and apply a similar method of rapid fire detection in a relatively small area of Chernobyl Exclusion Zone.

3. The purpose and objectives of this study

The purpose of this study is to analyze the methods of fire identification using low-resolution space images obtained from Terra Modis and NOAA satellites and determine the directions of information technology development. Consider Methods for detecting fires using low-resolution space images. There are a number of algorithms available to identify potentially fire pixels (PF). They are based on estimates of bright temperature in spectral ranges from 3.5–4 to 10.5–11.5 microns.

These algorithms are divided into – threshold and spatial (contextual) groups. Threshold algorithms are sensitive to conditions of a particular terrain, time of year, time of day, type of vegetation, soil temperature, etc. Contextual algorithms are based on comparing the brightness temperature of a potentially fire pixel with bright ness temperature cloudless surrounding pixels. This algorithm takes into account the spatial change of the thermal background, as opposed to the single-band or multiband threshold separation algorithm.

One of the problematic aspects in the Fire Detection Method using low-resolution space images is "Cloud and Water Masking". To identify "fire pixels", it is important to exclude from the analysis fragments of images that are covered with clouds and occupied by water objects.

It is important to divide the analysis methods in the daytime and at night. In the daytime, pixels are considered dimmed if the following condition is true: $(\rho_{0.65} + \rho_{0.85} > 0.9)$ OR $(T_{12} < 265\,\text{K})$ OR $(\rho_{0.65} + \rho_{0.85} > 0.7)$ AND $(T_{12} < 285\,\text{K})$. Where T_{12} is the brightness temperature in the spectral range of 11–12 microns, and $\rho_{0.65}$ and $\rho_{0.85}$ are reflective coefficients. At night, pixels were marked as cloud when the condition $T_{12} < 265$ K was met. It found that these simple criteria were sufficient to identify larger, cooler clouds, but did not take into account small clouds and cloud edges.

Algorithm for preliminary detection of potentially "fire" pixels for Modis images

Identification of pixels in which one or more fires are actively burning at the time of passing the satellite above the Earth is the basis of the algorithm for detecting potentially "fire" pixels: $T_4 > 310$ K, $\Delta T > 10$ K, $\rho_{0.86} < 0.3$, where $\Delta T = T_4 - T_{11}$, where T_4 is a bright spectral range of 4 μk, T_{11} is a bright temperature in the spectral range of 11 μk , $\rho_{0.86}$ is the reflection coefficient of the sun's um. For night pixels, the reflection test is skipped and the T_4 threshold has been reduced to 305 K. Pixels that do not match these previous tests are immediately classified as non-fire pixels. There are two logical paths, through which fire pixels can be identified. The first consists of a simple absolute threshold test. This threshold must be set high enough to trigger only very unambiguous fire pixels, that is, those that cannot become false alarms. The second way consists of a series of developed contextual tests to detect most active fire pixels that are less.

Absolute threshold test

The absolute threshold criterion remains identical to the original algorithm[4]: $T_4 > 360$ K (> 320 K at night). Despite the high daytime threshold, the usefulness of this test depends on the angle of reflection of sunlight above the horizon, otherwise false alarms caused by the glare of certain objects may occur.

Background algorithm

The next step of the algorithm, which is performed regardless of the result of the absolute threshold test, is an attempt to use neighboring pixels to assess the radiometric signal of a potential "fire" pixel. Real pixels are used as background values in the window, decentralized is a pixel identified as potentially "fire". In this window, pixels that are considered to be valid are considered to be:

1) contain useful observations;
2) are located on land;
3) not contaminated with clouds;
4) are not background "fire" pixels.

"Fire" pixels are defined as having $T_4 > 325$ K and $\Delta T > 20$ K for day observations, or $T_4 > 310$ K and $\Delta T > 10$ K for night observations

An important factor in this algorithm is the size of the window. Empirically expedient is a size of 21×21 pixels, with validity of at least 8 pixels.

The number of actual pixels in the background window is N_v. During the analysis process, the number of adjacent pixels (N_f) is determined, as well as the number of adjacent pixels excluded as water (N_w).

If a sufficient number of real adjacent pixels is identified, several statistical indicators are calculated statistical indicators. This is $\overline{T_4}$ and σ_4, the corresponding mean and the average absolute deviation of T_4 for real adjacent pixels. $\overline{T_{11}}$ and σ_{11} corresponding mean and absolute deviation of T_{11} for real adjacent pixels; $\overline{\Delta T}$ and $\sigma\Delta T$ corresponding to the mean and average for real adjacent pixels.

For contextual fire detection algorithms, the average absolute deviation is used as a measure of dispersion rather than standard deviation, since it is more resistant to deviations. This is very desirable, since pollution of the background window with uncertain clouds, water, fires and other sources is not uncommon.

Contextual algorithm for identifying fire pixels using NOAA AVHHR images

Contextual algorithm that identifies "fire" pixels results in two consecutive tests:

Test 1: pixel selected as "fire" if:

$T^B(3) > 311$ K [1],

$T^B(3) - T^B(4) > 8$ K [2],

Test 2: The pixel is confirmed as "fire", if:

$T^B(3\text{-}4)_{PF} - [T^B(3\text{-}4)_b + \delta T^B(3\text{-}4)_b] > 0$ [3],

$T^B(3)_{PF} - [T^B(3)_b + \delta T^B(3)_b] > 3K$ [4],

where $T^B(3)_b$ – the average value of the bright background temperature in band 3, $\delta T^B(3)_b$ – standard deviation of the bright background temperature in the band 3, $T^B(3\text{-}4)_b$ – the average value of the difference in bright temperature (band 3 – band 4), $\delta T(3\text{-}4)_b$ – standard deviation of the difference in bright temperature (band 3 – band 4).

The pixels selected in this way can be divided into three groups: definitely firefighters, probably firefighters and unambiguously non-fire. Further determination of unambiguously fire pixels is carried out by masking places where PF is uniquely inseparable (sandy soils, reservoirs, man-made objects that emit a large amount of heat).

The presence of clearly defined areas with a high degree of fire danger (maps of forest quarters in the forest areas of the exclusion zone with certain characteristics of the state of forests) makes it possible to attribute PF do to unambiguous-fire pixels with a high probability.

Implementation of the fire identification algorithm. The method was implemented with the help of the Visual Programming Tool PowerBuilder in the data processing system of remote sensing Erdas Imaging. The algorithm of implementation of this method is shown in Figure 1.

Fire monitoring should be carried out by a sequence of operations in the environment of the ERDAS remote sensing system. This sequence can be divided into several stages:

1) acquisition and adaptation of space imagery
2) PF identification according to criteria 1-
3) geometric correction of the image
4) integration of the snapshot with cartographic data and interpretation of the results.

Reception stations receive NOAA AVHRR pictures in HRPT format. Specialized program SmartTrack AVHRR snapshot is converted to intermediate Level1B format, and Import ErdasImagine module intermediate file is converted to the working format of the system ErdasImagine.

During import, radiometric correction (data calibration) is performed. The result of this processing stage is a fragment of an imported image covering the territory of Ukraine and adjacent regions of other countries (Figure 2).

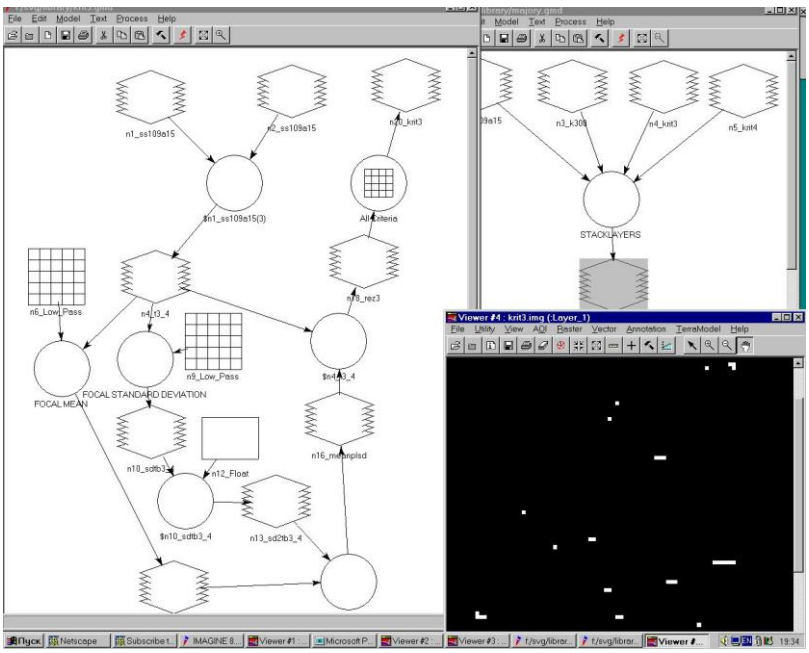


Fig. 1. Technology of automatic identification of fires

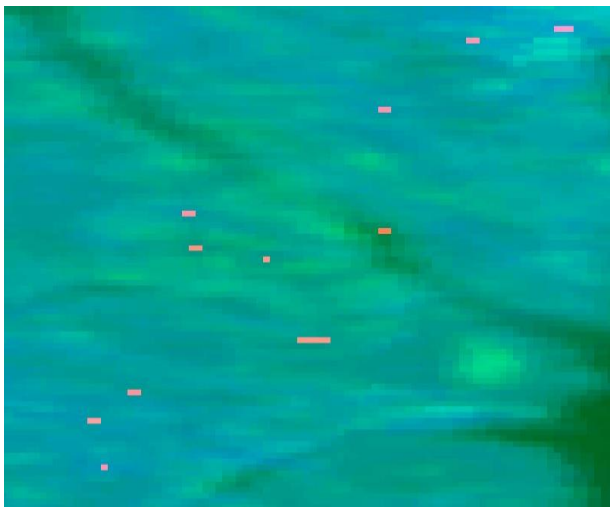


Fig. 2. Potentially fire pixels

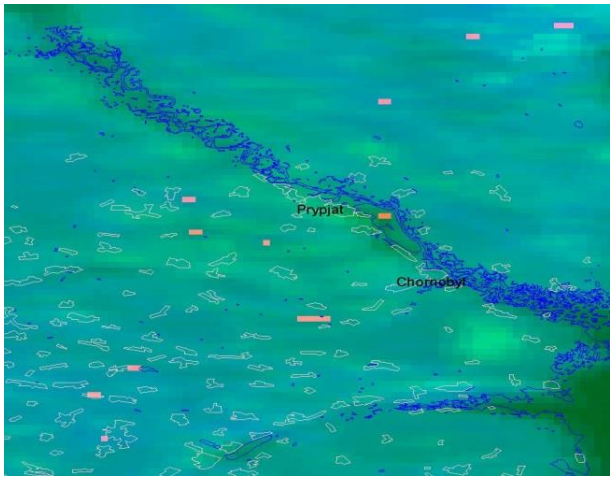


Fig. 3. Rectified image is combined with a vector layer of reservoirs

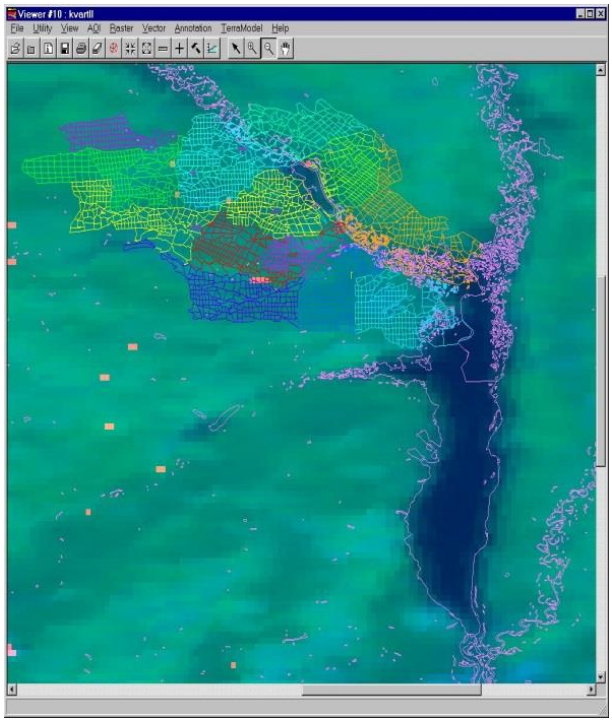


Fig. 4. Identification of the potentially fire pixels of the summaries from the map of the forest quarters

For PF identification as part of the model project (ModelMaker package procedures – Erdas Imagine module), that implement the criteria 1–3, and so-as-cloud masking models (deleting from analysis area)

The process of allocating fires has the following sequence of actions: Selecting pixels according to criterion [1] – selection of pixels with a bright > 311 K.

The selection of pixels according to the criterion [2] – selection of pixels by temperature difference in band 3 and 4. Potentially fire pixels highlighted in the second stage appear bright red in the picture (Figure 3). The selection of pixels according to the criterion [3] – statistical assessment of the difference between brightness temperatures in band 3 and 4. Selection of pixels according to the criterion [4] – statistical assessment of the temperature in band 3.

The result of these four stages will be pictures on which potentially fire pixels will be allocated. The intersection of all four criteria shows us with a fairly high probability of having fires in the area under study. As a result, we will get a picture with the identified fires.

Geometric correction of the image After identifying fires, it is necessary to determine the exact localization (in this project, the binding with accuracy to the forest quarter is sufficient). To do this, the picture must be corrected (rectified) and brought to the cartographic projection, which contains the main cartographic data associated with the Chornobyl exclusion zone.

This process was carried out by identifying the anchor points in the picture and matching them with objects on a vector map Geometric correction was carried out on the basis of the method of polynomial transformation of the second order. Figure 4 shows the following picture. The quality of geometric correction is visible in combination with the vector layer of water surfaces.

Obtaining information based on the processing of SES data during May and June, the above-mentioned method was tested on fragments of images of the exclusion zone. Composite time series of snapshot fragments. On the basis of this technique, several constantly "hot" objects were found in the zone (these are heated sandy soils near the Uzh River, the RV storage facility near Buriakivka village and a number of others.

4. Conclusions

The problem of fires in recent years attracts special attention in the context of the growing influence of such global processes as the reduction of the world's forest areas, the loss of biodiversity, global climate change and land use changes. This is due to the complexity and ambiguity of the impact of fires on forests, environments and communities living around forests. Determining the areas of forests covered by fires is an important component of the information support of forestry. The method proposed in this work, which is based on a combination of medium infrared (3–4 microns) with thermal (10–11 micron) NOAA AVHRR space images, allows you to quickly determine the fire in real time at the sub-pixel level. Prompt detection of fires makes it possible to quickly extinguish fires and save forest resources. An example of this method was carried out on the fires of the Chornobyl Exclusion Zone and fires at gas wells in the Poltava region of Ukraine.

To quickly determine the fire in real time at the subpixel level. Prompt detection of fires makes it possible to quickly extinguish fires and save forest resources. An example of this method was carried out on the fires of the Chornobyl Exclusion Zone and fires at gas wells in the Poltava region of Ukraine.

This method is universal and requires only high-quality NOAA AVHRR space images, Terra Modis, which is freely available.

References

[1] Chandra A. M., Gosh S. K.: Remote sensing and geographic information systems. Tecnosfera, Moscow 2008.
[2] Griffiths P. et al.: Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. Remote Sensing of Environment 151/2014, 72–88.
[3] Justice C. O. et al.: Robust statistics. Wiley, New York 2002.
[4] Kaufman Y. J., Justice C. O., Flynn L. P., Kendall J. D., Prins E. M., Giglio L., Ward D. E., Menzel W. P., Setzer A. W.: Potential global fire monitoring from EOS-MODIS. Journal of Geophysical Research 103(D24)/1998, 32215–32238.
[5] Kashkin V. B., Suchinin A. I.: Remote sensing of the Earth from space. Digital imaging. Logos, Moscow 2001.
[6] Krycuk S. G.: Mapping boreal forests using satellite data. Modern problems of remote sensing of the Earth from space 9(4)/2012, 255–264.
[7] Morisette J. T. et al.: Identification of subresolution high temperature sources using a thermal IR sensor. Photogrammetric Engineering and Remote Sensing 47/2003, 1311–1318.
[8] Morissette, J. T. et al.: Validation of the MODIS active fire product over Southern Africa with ASTER data. International Journal of Remote Sensing 26(19)/2002, 4239–4264.
[9] Seielstad C. A. et al.: MODIS level 1A Earth location: Algorithm theoretical basis document version 3.0. SDST-092, MODIS Science Data Support Team 2002.
[10] Stroppiana D. et al.: Testing the sensitivity of a MODIS-like daytime active fire detection model in Alaska using NOAA/AVHRR infrared data. Photogrammetric Engineering and Remote Sensing 68/2000, 831–838.
[11] Strugailo V. V.: Review of methods of filtering and segmentation of digital images. Science and education 5/2012, 270–281.
[12] Vyshnjakov V. J., Tkachuk P. A.: Features of methods for determining temperature anomalies according to remote sensing MODIS (TERRA) and AVHRR (NOAA). Estimates of their quality. Ecological safety and nature management 10/2012, 81–90.
[13] Zibcev S. V., Myronjuk V. V.: Methodical recommendations from MODIS and Landsat for monitoring of forest fires: Recommendations for forestry enterprises. NUBandN, Kiev 2015.

M.Sc. Valerii Shvaiko
e-mail: valshvaiko57@gmail.com

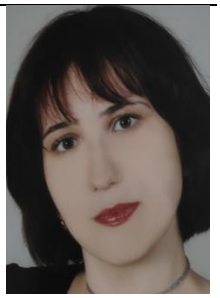
Valerii Shvaiko is currently lecturer Department Automation of Design of Energy Processes and Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. Research interests – geoinformation system, remote sensing, civil protection of the population.

http://orcid.org/0000-0002-9304-8710

M.Sc. Olena Bandurka
e-mail: o.i.bandurka@ukr.net

Olena Bandurka is currently postgraduate student the Department Automation of Design of Energy Processes and Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. Research interests – knowledge-based geospatial analysis techniques for prediction and assessment of impact on biological environment.

http://orcid.org/0000-0002-8059-1861



Ph.D. Vadym Shpuryk
e-mail: wadimoff@gmail.com

Ph.D. doctoral supervisor, associate professor of the Department of Computer Aided Design at National Technical University, Ukraine. Engaged in the research of artificial intelligence, GIS technology, underwater sonar signal processing and data processing algorithms. A dedicated system professional and software developer with 25+ years of technology development and management experience. Programmer/Analyst with solid background in all phases of software development life cycle of various applications.

http://orcid.org/0000-0002-3477-5731

Prof. Yevhen V. Havrylko
e-mail: gev.1964@ukr.net

Doctor of Engineering Sciences, Professor, Professor department Automation of Design of Energy Processes and Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.
Research interests – ecological safety, civil protection of the population, remote sensing of the Land, systems of satellite telecommunications.

http://orcid.org/0000-0001-9437-3964

otrzymano/received: 27.02.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2579

GENERATING FIRE-PROOF CURTAINS BY EXPLOSION-PRODUCTION OF WATER AEROSOL AS AN ELEMENT OF FIRE-SAFETY ENGINEERING

Grzegorz Śmigielski
Kazimierz Wielki University, Institute of Computer Science, Bydgoszcz, Poland

Abstract. *Parameters of a fire-quenching system based on explosive water-aerosol production and spreading are presented. Such a system with correctly tailored amounts of the water and the explosive material produces water-aerosol with high fire-quenching efficiency due to the small radii of the droplets and at the same time secures that shock-wave pressures are safe for the human body at distances exceeding 30 m from the explosion axis.*

Keywords: explosive aerosol production, firefighting

WYTWARZANIE ZAPÓR OGNIOWYCH REALIZOWANYCH POPRZEZ WYBUCHOWE WYTWARZANIE AEROZOLU WODNEGO JAKO ELEMENT INŻYNIERII BEZPIECZEŃSTWA POŻAROWEGO

Streszczenie. *Przedstawiono parametry instalacji gaśniczej opartej na wytwarzaniu i rozprzestrzenianiu się wybuchowego aerozolu wodnego. Taki system z odpowiednio dobranymi ilościami wody i materiału wybuchowego wytwarza aerozol wodny o dużej skuteczności gaszenia pożaru dzięki małym promieniom kropelek i jednocześnie zapewnia, że ciśnienie fali uderzeniowej jest bezpieczne dla ludzi w odległości powyżej 30 m od osi wybuchu.*

Słowa kluczowe: wybuchowe wytwarzanie aerozolu, gaszenie pożarów

Introduction

Production of water-aerosol with explosives is a new technology. Though the idea of using water-aerosol in fire-fighting and some its practical applications are widely known [9, 19, 22]. The novelty consists in the way of using the aerosol and in the scale of its production. Explosion-produced water-aerosol may be used for extinguishing large-scale fires provided it is produced in sufficiently large amount. Research on practical implementation of the idea of producing aerosol with explosive is being carried on for several years [4–7, 15–18]. The results obtained till now are very promising since high efficiency of this solution was achieved – fire of inflamed wooden boards, covering a disc of 30 m in diameter had been quenched with about 1000 dm^3 of water.

The method of producing aerosol by explosion may be used not only for fire extinguishing. There are many other potential applications and proliferation of the method and its implementation into industrial production will entail appearance of other applications. Such aerosol can be used for neutralization of chemically and/or radioactively contaminated areas, spraying vegetation with chemicals or deposition of munitions. Usage of the method for such diverse application requires only proper selection of the size of the bag and the kind of the liquid of which the aerosol is to be produced.

The explosive production of water-aerosol is in a number of ways related to the safety engineering. First, high fire-quenching efficiency of the method recommends it for using in generating fire-proof curtains. Second, the method generates some safety problems, and consequently, it is necessary to determine safety conditions for its applications. In particular it is necessary to determine the area around the explosion point in which shock wave pressure may be dangerous for people. Separate safety problems are connected with the process of production and storage of the water bags and fixing explosive charges inside them.

1. Production of water-aerosol

This publication is based on the idea of producing water-aerosol by explosive method. It consists in detonating explosive charge placed inside a water container. The detonation causes fragmentation (pulverization) of water into droplets that move symmetrically away from the explosion axis. The aerosol generated this way covers certain area which in particular may be in fire.

Water capsule, i.e., the container with water and explosive material (Fig. 1 for its schematic cross-section and Fig. 2 for its actual view) secures fixed position of the latter with respect to the contained bulk of water. The water – explosive system is expected to optimize production of water-aerosol cloud with droplets of small enough diameters and covering densely enough a desired area. In addition, the capsule's construction must fulfill safety requirements [21].

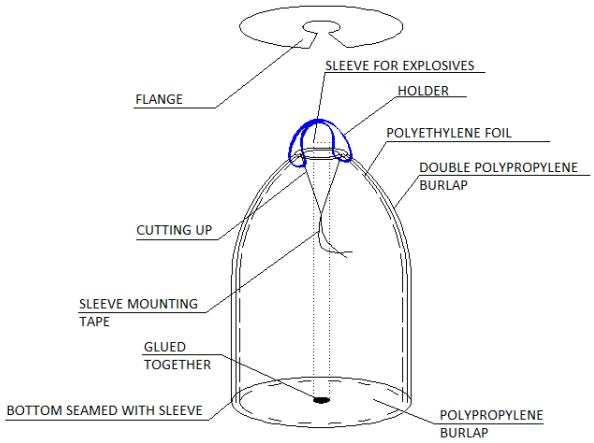


Fig. 1. Water capsule worked-out during field tests – scheme of the design

Fig. 2. Water capsule worked-out during field tests – the outer view of an actual capsule

Submerging explosive in the water meets the above two requirements to a large degree, since it secures high efficiency of the explosion energy transfer to the bulk of water and quenching of the shock-way to the level safe for the humans even at moderate distances from the explosion axis [13]. The production cost, weight and suitability for storage and transportation are other factors that have to be taken into account while the commercial construction is considered.

2. Parameters of water aerosol as fire-extinguishing agent

The explosion ensues fragmentation of water contained in the capsule and acquiring certain initial axifugal velocity by the produced aerosol droplets. After several milliseconds from the explosion the droplets cease to accelerate and start to move with deceleration caused by the drag and friction caused by the air [7]. The drag force during the motion with large velocity is considerable and causes further fragmentation (called "stripping") of droplets [7, 12, 14]. This process continues till the size of droplets and their velocities are reduced to such values that the drag force is to small to cause further fragmentation. Therefore the sizes of aerosol droplets change in the course of cloud's expansion, and from the point of view of applications of explosive aerosoling it is the final droplet size that counts.

Explosive production of aerosol required some investigation of the prospects of practical application of the method. Among others it was necessary to measure: pressure of the shock-wave generated by the explosion, cloud's diameter and size of droplets. The research stand was shown in Fig. 3. and Fig. 4.

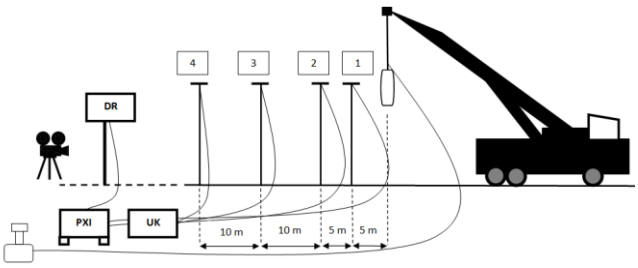


Fig. 3. Scheme of the research stand

Fig. 4. The research stand before capsule's explosion

2.1. Measurement of shock-wave pressure

The apparatus was composed of a short-circuit detector, triggering measurement, four piezoelectric high pressure sensors (IEPE) 137xx series [26], a system of conditioning the signal (UK) [24] and an industrial NI PXI computer [23]. The points of the sensors are directed towards the explosion axis to reduce the drag force and to prevent blowing them out from the poles or from damage by expanding shock-wave and explosion products. To ensure comparable relative accuracy of pressure measurements at all four distances from the explosion axis, three types of sensors are used and their parameters are given in Table 1. A schematic view of the configuration of the sensors (no. 1 – no. 4) with respect to the water capsule is shown in Fig. 3.

The measurement is triggered on at the moment of explosion of the charge by the upset sensor inserted into the earlier. The measuring systems allows to register 10^5 samples of pressure at the sampling frequency $5 \cdot 10^5$ samples per second. It allows one to cover time interval of 200 ms with time resolution 2 μs. Acquisition and processing of the results of measurements is made possible by an industrial PXI computer supplied with the software LabVIEW RT, and equipped with a measurement card. Time-profiles of the shock-wave pressure registered by the measuring system allowed one to determine pressure distribution with respect to the distance from the explosion axis.

Table 1. Parameters of the pressure sensors

Number of sensor	Sensitivity K_1 [mV/kPa]	Maximum pressure [MPa]	Resolution [kPa]	Expanded uncertainty $U_r(K_1)$ [%] (95%)
1	0.149	34.5	0.69	1.3
2	2.8	34.5	0.001	0.8
3	15.1	6.9	0.069	0.8
4	13.7	6.9	0.069	1.3

The maximum pressures registered by the sensors for various conditions are shown in Fig. 5 – Fig. 8. The tests were conducted for three various capsule capacities (600, 1200 and 1500 dm^3 – size of the bubbles), three various types of explosives (EMULINIT [25] – red bubbles, Plastic C4 – yellow, Saletrol – blue) and a number of explosion energies.

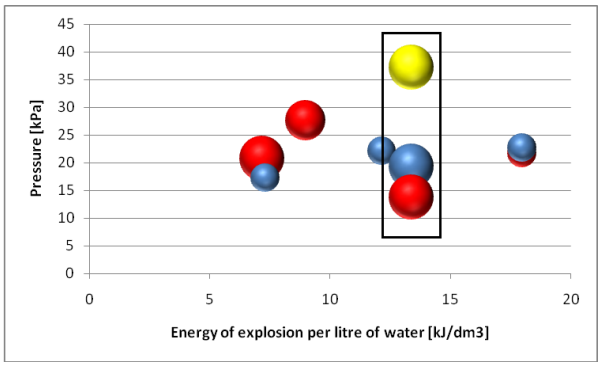


Fig. 5. Maximum pressure registered by sensor no. 1

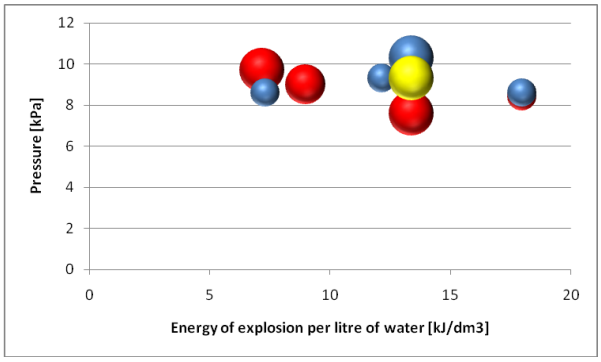


Fig. 6. Maximum pressure registered by sensor no. 2

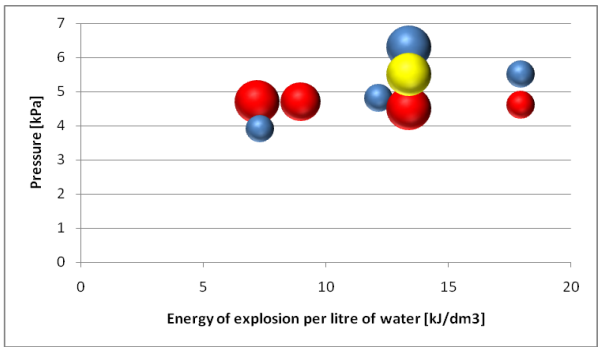


Fig. 7. Maximum pressure registered by sensor no. 3

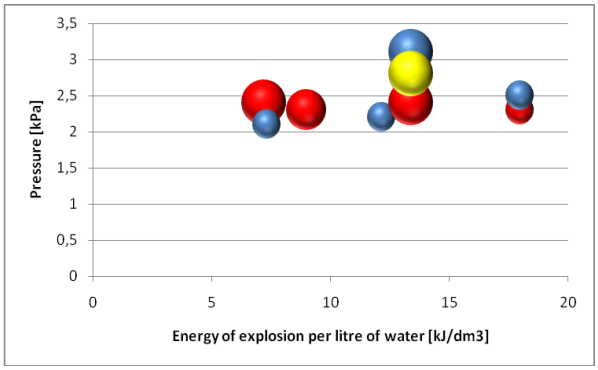


Fig. 8. Maximum pressure registered by sensor no. 4

Assuming that the pressure amplitude of the shock-wave front at a distance from the explosion axis is a measure of the "idle" energy, i.e. the energy that was not converted into the useful work over the water-bulk, the largest maximum pressure on the sensor No. 1 would correspond to the least efficient configuration and the smallest maximum to the most efficient configuration.

Looking specifically at the rectangle mark in Fig. 5 one can notice that for the same amount of water – 1500 kg, and the same explosion energy – 20.1 MJ, which corresponds to 13.4 kJ/dm^3, various kinds of tested explosive materials exhibit various efficiencies. Apparently from this point of view Plastic (C4) seems to be the least efficient while EMULINIT seems to be the most efficient of the three explosive material tested.

The pressure in the shock-wave front 30 m from the explosion axis in no test (independent of the capsule capacity and explosion energy) exceeded 3 kPa [7], i.e. about 3% of the normal atmospheric pressure, which is safe for the human body [11, 21].

The pressure measurements uncertainty is connected with errors of sensors and errors of measurement chain – signal conditioning system and input analog card.

The value of the pressure can be determined using the following relation [17]:

$$p = \frac{U_2}{K_1 K_2}, \tag{1}$$

where K_1 is a coefficient of the pressure/voltage processing, U_2 is measured voltage and K_2 is amplification of the signal conditioning system. The value of the sensors pressure measurements uncertainty has been determined based on the value of expanded uncertainty taken from calibration certificate (Table 1).

The overall uncertainty of the complex measurement has been evaluated based on the following expression [1, 17, 20]:

$$u(p) = \sqrt{\left(\frac{\partial p}{\partial K_1}\right)^2 u^2(K_1) + \left(\frac{\partial p}{\partial K_2}\right)^2 u^2(K_2) + \left(\frac{\partial p}{\partial U_2}\right)^2 u^2(U_2)}, \tag{2}$$

The maximum uncertainty values for individual sensors were: $u_1(p)$ = 0.22 kPa, $u_2(p)$ = 0.067 kPa, $u_3(p)$ = 0.026 kPa, $u_4(p)$ = 0.021 kPa.

2.2. Measurement of cloud's diameter

Measurements of cloud's diameter had to be performed indirectly. The method used during our experiments was based on registration of the process of expansion with a fast video camera working at the frequency 250 fps. The camera was placed at a distance from the axis of explosion much larger than the maximum cloud's diameter (Fig. 3, Fig. 4, Fig. 9). This way it was safe against possible influence by the shockwave and, in addition, its distant position of the helped to minimize parallax error.

The diameter of the cloud was estimated from subsequent frames by its comparison with the reference scale indicated either by stakes (Fig. 4). The results of measurements are presented in Table 2.

Analyzing a number of tests it has been found that the maximal error of the cloud diameter measurement can be associated with a cloud spreading time, for the time over 2 sec it is on Δx = 2 m [17].

In general, for bags of the same size increase of cloud's diameter is observed with increase explosion energy up to a limiting energy value that depends on the water-bag size. Using energies exceeding the limiting value results in quite a pronounced decrease of cloud's diameter, which can be interpreted in terms of the breakdown of efficiency of energy transfer to the water-bulk. In the case of energies below the limiting values the shock-wave can be perceived at distances larger than 50 m only acoustically. On the contrary, while the limiting energy value is exceeded, the shock-wave can be perceived directly by observers as a "blow", and can cause some damages to objects like cars even that far as 100 m from the explosion axis.

Fig. 9. The aerosol cloud a few dozen milliseconds after the explosion

Table 2. The maximum spray cloud diameter (Emulinit)

Capsule capacity	Explosion energy	Cloud diameter
dm^3	kJ	m
600	4384	24
600	7293	32
1200	7293	38
1200	13110	38
1200	16601	41
1500	10784	40
1500	13110	47
1500	16601	49
1500	20092	53

2.3. Measurement of droplets size

In this section results of measurements of explosion-produced aerosol droplets are presented. The method of measurement, called optical method [4, 18] applied droplet recorder – an apparatus located in the line of droplets' trajectories (Fig. 3 – DR).

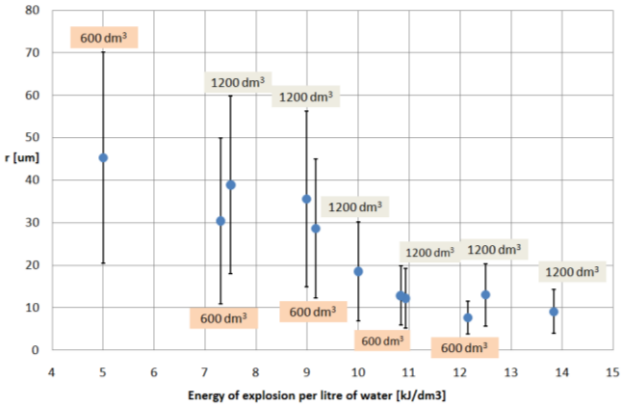


Fig. 10. Dependence on the explosion energy of the average droplet radius (approximating expectation value of their distribution) together with measurement uncertainty equal to the standard deviation

Determination of the radius distribution of droplets in this method was based on collecting them on a glass plate and measuring diameters of adhering droplets using microscope. The shutter is the core of the recorder. It opens with time delay (with respect to explosion) and for time interval chosen by the operator under the control of the recorder driver [4, 18].

The optical method allows one to measure sizes of aerosol droplets inside the cloud. In Fig. 10. a plot of average droplet radius versus the explosion energy per 1 litre of sprayed liquid is shown. The average radii range from about 7 μm to about 45 μm.

3. Summary

Water aerosol is a perfect fire-extinguishing agent whose fire quenching efficiency surpasses many times that of jets of water. In the case of explosive aerosoling the fire-extinguishing effect is enhanced by the shock-wave. The fire-extinguishing properties can be further enhanced and modified by adding various admixtures to the basic medium, which can considerably extend the range of its application.

The research performed for three different types of the explosive charge: Saletrol (ANFO), Emulinit and plastic explosive (C4) showed that the Emulinit is most efficient in creation of the aerosol cloud.

Basic parameters of the water-capsule, which have been established are shown in Table 3. The amount of the explosive material used for the production of aerosol is chosen depending on the size of the capsule to secure reduction of the shock-wave pressure amplitude to the level secure for human body at distances from the axis of explosion larger than 30 m [13, 21].

Table 3.Basic parameters of water-capsule

Capacity of water-capsule	Bottom diameter	Container height	Amount of the main explosive charge (Emulinit 2)			Initiating charge ERGODYN 37 SE
			Ver. A	Ver. B	Ver. C	
dm^3	mm	mm	dag			g
600	900	1250	125	250	400	500
1200	900	1650	400	500	650	500
1800	900	2200	500	650	800	500

Generation of fire-proof aerosol curtains remains in the area of interest of the safety engineering. Applying water-capsules for production of such curtains seems to be a reasonable idea. All experiments and measurements performed by us provided evidences in favor of large potential of such a solution. Further research will certainly help to exploit this potential to a larger degree. A common danger caused by fires should be a sufficient justification for further research. Novelty of our research consists not that much in the idea of explosive generation of aerosol as in the scale of such production. Production of water-aerosol on the scale discussed in this publication was not reported by any other research group, although aerosoling of amounts of liquid smaller by three orders of magnitude was investigated [2, 3, 10, 12]. That, however, was not sufficient from the point of view of practical applications.

Acknowledgments

The author is grateful to Prof. Roman Dygdała and Dr. Eng. Damian Lewandowski for his assistance and collaboration during the tests on military training area and analysis of the results.

References

[1] Arendarski J.: Niepewność pomiarów. OWPW, Warszawa 2006.
[2] Brenguier J., Pawlowska H., Schüller L., Preusker R., Fischer J., Fouquart Y.: Radiative Properties of Boundary Layer Clouds: Droplet Effective Radius Versus Number Concentration. J. Atm. Sci. 57/2000, 803–821.
[3] Breon F., Goloub P.: Cloud Droplet Effective Radius From Space Borne Polarization Measurements. Geophys. Res. Lett. 25/1998, 1879–1882.
[4] Chaberski D., Grzelak S., Lewandowski D., Dygdała R., Zieliński M., Stefański K., Śmigielski G.: Measuring Distribution of Radii of Droplets Forming Explosively Generated Water-Spray Cloud. Metrol. Meas. Syst. 3/2010, 363–382.
[5] Dygdała R., Stefański K., Śmigielski G., Lewandowski D., Kaczorowski M.: Aerosol Produced by Explosive Detonation. Pomiary, Automatyka, Kontrola 53(9)/2007, 357–360.
[6] Kuhrt E., Knollenberg J., Martens V.: An automatic early warning system for forest fires. Annals of Burns and Fire Disasters 14/2001, 151.
[7] Lewandowski D.: Wytwarzanie aerozolu wodnego metodą wybuchową – aspekty metrologiczne, Praca doktorska. Politechnika Wrocławska. Wrocław 2011.
[8] Lewicki J.: Prognozowanie wielkości zagrożeń powstałych przy prowadzeniu robót strzałowych w budownictwie. Górnictwo i Geoinżynieria 28(3/1)/2004, 251–267.
[9] Liu Z., Kim A. K., Carpenter D.: Extinguishment of large cooking oil pool fires by the use of water mist systems. Combustion Institute/Canada Section, Spring Technical Meeting, 2004, 1–6.
[10] Morka A. Wybuchowe rozpraszanie cieczy, PWN, Warszawa 2001.
[11] Onderka Z., Sieradzki J., Winzer J.: Wpływ robót strzelniczych na otoczenie kopalń odkrywkowych. UWND AGH, Kraków 2003.
[12] Papliński A.: Modelowanie spalania i wybuchu w niejednorodnych fizycznie ośrodkach reaktywnych. Wojskowa Akademia Techniczna, Warszawa 2009.
[13] Rosenkiewicz D., Bogdański M., Zarzycki J., Łabędzki J.: Analiza parametrów fal podmuchowych generowanych ładunkami gaśniczymi. II Międzynarodowa Konferencja IPOEX 2005: Materiały wybuchowe, Badania – Zastosowanie – Bezpieczeństwo, Tom I, 2006.
[14] Stebnovskii S. V.: Pulsed Dispersion as the Critical Regime of Destruction of a Liquid. Combustion, Explosion, and Shock Waves 44(2)/2008, 228–238.
[15] Stefański K., Lewandowski D., Dygdała R., Kaczorowski M., Ingwer-Żabowska M., Śmigielski G., Papliński A.: Explosive formation and spreading of water-spray cloud – experimental development and model analyses. Central European Journal of Energetic Materials 6(3-4)/2009, 291–302.
[16] Śmigielski G., Toczek W., Dygdała R., Stefański K.: Metrological Analysis of Precision of The System of Delivering Water-Capsule for Explosive Production of Water Aerosol. Metrol. Meas. Syst. XXIII(1)/2016, 47–58.
[17] Śmigielski G., Dygdała R., Kaczorowski M., Serejko G.: Measurements of Parameters of Water Aerosol Obtained by Explosive Method. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOŚ 3/2016, 32–35.
[18] Śmigielski G. et. al.: Rejestrator kropel aerozolu wodnego wytwarzanego metodą wybuchową. PAK 60(12)/2014, 1170–1173.
[19] Teie W.: Firefighter's Handbook of Wildland Firefighting. Dear Valley Press, Recue 1994.
[20] Zięba A.: Analiza danych w naukach ścisłych i technice. PWN, Warszawa 2013.
[21] Dz. U. Nr 163, poz. 1577 – Rozporządzenie ministra gospodarki, pracy i polityki społecznej w sprawie bezpieczeństwa i higieny pracy przy produkcji, transporcie wewnątrzzakładowym oraz obrocie materiałów wybuchowych, w tym wyrobów pirotechnicznych.
[22] http://www.telesto.pl
[23] http://www.ni.com
[24] http://www.vibx.pl
[25] http://www.nitroerg.pl
[26] http://www.pcb.com

Ph.D. Grzegorz Śmigielski
e-mail: gsmigielski@ukw.edu.pl

Graduate of Faculty of Physics, Astronomy and Applied Informatics Nicolaus Copernicus University in Toruń. He received the Ph.D. degree in electronics from Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics in 2011. His main scientific interests are control and measurement systems.

http://orcid.org/0000-0003-3781-0894

otrzymano/received: 28.02.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2430

METHODS FOR ASSESSMENT AND FORECASTING OF ELECTROMAGNETIC RADIATION LEVELS IN URBAN ENVIRONMENTS

Denys Bakhtiiarov, Oleksandr Lavrynenko, Nataliia Lishchynovska, Ivan Basiuk, Tetiana Prykhodko
National Aviation University, Kiev, Ukraine

Abstract. *Methods for synthesis a structural diagram of the processes for detecting and locating technical information leakage channels are analysed. Software for defining a controlled room zone was also presented. A proprietary approach to search the electromagnetic environment under radio interference has been developed to detect devices for unauthorized control of acoustic information using microphones and transmission of this information using a radio channel.*

Keywords: side electromagnetic radiation, controlled zone, radio wave propagation, electromagnetic environment monitoring

METODY OCENY I PROGNOZOWANIA POZIOMÓW PROMIENIOWANIA ELEKTROMAGNETYCZNEGO W ŚRODOWISKACH MIEJSKICH

Streszczenie. *Przedstawiono metody syntezy schematu strukturalnego procesów wykrywania i lokalizacji technicznych kanałów wycieku informacji. Przedstawiono także oprogramowanie do definiowania kontrolowanej strefy pomieszczenia. Opracowano autorskie podejście do przeszukiwania środowiska elektromagnetycznego w warunkach zakłóceń radiowych, w celu wykrywania urządzeń służących do niedozwolonej kontroli informacji akustycznej za pomocą mikrofonów i transmisji tej informacji z wykorzystaniem kanału radiowego.*

Słowa kluczowe: niepożądane promieniowanie elektromagnetyczne, strefa kontrolowana, propagacja fal radiowych, monitoring środowiska elektromagnetycznego

Introduction

One of the areas of monitoring the electromagnetic environment is the search and detection of specially organized and potential radio channels for information leakage [9, 12].

The duration of the search process and the reliability of the received information depend on the completeness of solving tasks by search equipment. The completeness and speed of their implementation, the effectiveness of the search engine, the reliability of the information obtained and the probability of decision-making depend on the structure of the search engine and the characteristics of the tools used in it [6, 8].

1. Detection of technical leakage channels

Detection of technical leakage channels is a complex multi-stage process, which in a simplified form can be represented as a set of a number of stages presented in Fig. 1.

And the **Stage I** involves the analysis of the current load range and the accumulation of data on the frequencies, levels and nature of electromagnetic radiation in the operating frequency range with the binding of data to the receiving location. "Known" radiation means a set of data accumulated over a certain period of time on the loading of the range, obtained by the results of the current control. It is assumed that there are no dangerous signals, which is achieved, for example, by the gradual accumulation of "known" radiation with careful verification of each of the radiation.

The list of "unknowns" (**Stage II**) includes data on radiation, the set of parameters which meet the specified search criteria. The use of a "reference" antenna involves the presence in the search engine of an antenna switch that provides alternate connection of one of the receiving (in a dedicated room) antennas and a "reference" antenna that is outside the controlled area, but provides reliable reception of all external signals.

Stage III involves testing, which gives a certain effect both when detecting radio microphones without closing (radiation in a dedicated room specially synthesized acoustic signals) and when conducting special studies on IEMRI by appropriate modulation of informative radiation parameters [2].

To perform **Stage IV**, it is necessary to compare the maximum (from the outputs of antennas in the controlled area) components of the spectrum with the levels of the corresponding components previously accumulated in the selected room "known" electromagnetic radiation (in the absence of radiation) and the limit level comparison, the decision is made on the presence (absence) of "unknown" radiation in the controlled area.

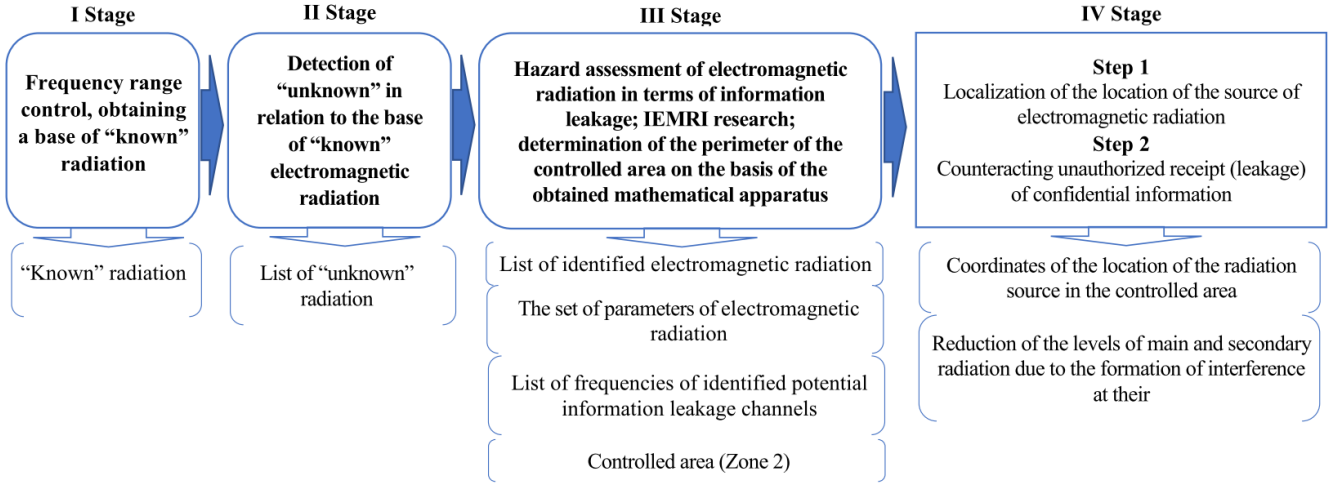


Fig. 1. The main stages of the process of identifying technical channels of information leakage

The coordinates for monitoring the electromagnetic environment are selected from the set of spectral samples $X_R\,(J,\,n)$ averaged over R realizations of the energy spectrum.

$$X_R(j,n)=\frac{1}{R}\times\sum_{r=1}^{R}\left|c_{(r)}(n)\right|^2 \tag{1}$$

where j is the number of the antenna connected to the input of the monitoring equipment.

The average power of a random process $U_m(t)$, represented in the frequency domain by a set of values of $X_R(j,n)$ numbers from n_{min} to n_{max}, is proportional to the sum of these samples [4]:

$$\hat{P}_{j,m}=10\lg\left(\sum_{n=n_{min}}^{n_{max}}X_R(j,n)\right)+\mu, \tag{2}$$

where μ – correction factor, which is determined by calibration of the antenna and equipment of the used channel for monitoring the electromagnetic environment (MEE).

The proposed structure of the search engine, which implements this algorithm and provides increased integrated sensitivity and maximum speed, contains: a set of wide-range antennas, one of which ("reference") is made outside the dedicated room; controlled antenna switch; frequency-controlled receiving path with IF bandwidth; analog-to-digital processing device based on fast Fourier transform; control device with a variable structure, which is determined by the proposed software.

Improving the efficiency of the use of MEE equipment in this method is provided by: the use of panoramic analysis based on FFT; reducing the amount of data processed when using a database of "known" electromagnetic radiation or signals from the output of the "reference" antenna.

This method is based on the known position of electrodynamics about the different nature of changes in the electromagnetic field strength in the near and far zones [1, 4, 5, 10, 11]. The components of the electric field vector emitted by the electric dipole p in spherical coordinates are determined by the expressions:

$$\begin{cases}E_r=\dfrac{1}{2\pi\varepsilon}\left(\dfrac{1}{r^3}-\dfrac{ik}{r^2}\right)\cos\Theta\left|\vec{p}\right|\exp\left[-i\omega t\right]\\ E_\Theta=\dfrac{1}{4\pi\varepsilon}\left(\dfrac{1}{r^3}-\dfrac{ik}{r^2}-\dfrac{k^2}{r}\right)\sin\Theta\left|\vec{p}\right|\exp\left[-i\omega t\right]\\ E_\varphi=0\end{cases} \tag{3}$$

where r, Θ, φ – spherical coordinates, E_r, E_θ, E_φ – components of the electric field in spherical coordinates, ε, μ – electric and magnetic permeability of free medium, ω – angular frequency of radiation.

The dipole moment is related to the radiated power W by the ratio:

$$W=\frac{\omega^4}{12\pi}\mu\sqrt{\varepsilon\mu}\left|\vec{p}\right|^2, \tag{4}$$

The module of electric field strength is determined by the expression [11]:

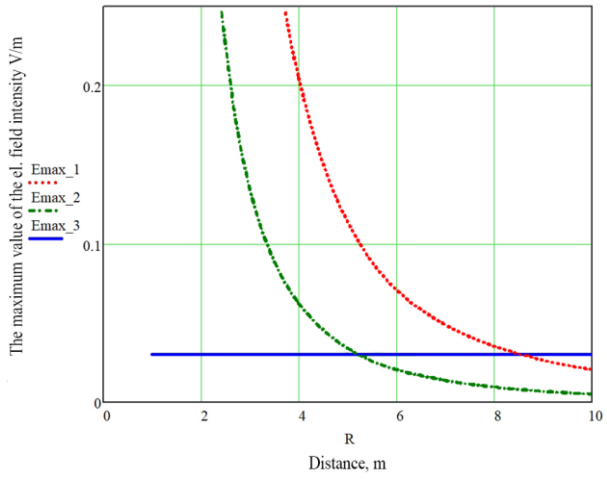
$$\left|\vec{E}\right|=\sqrt{\frac{12\pi W}{\mu\sqrt{\varepsilon\mu\omega^4}}}\cdot\left(\frac{1}{(2\pi\varepsilon)^2}\left(\frac{1}{r6}+\frac{k^2}{r^4}\right)\cos^2\Theta+\frac{1}{(4\pi\varepsilon)^2}\left(\left(\frac{1}{r^3}+\frac{k^2}{r}\right)^2+\frac{k^2}{r^4}\right)\sin^2\Theta\right)^{\frac{1}{2}} \tag{5}$$

The maximum modulus of electric field strength in all possible directions Θ, for a given r is determined by the expression [11]:

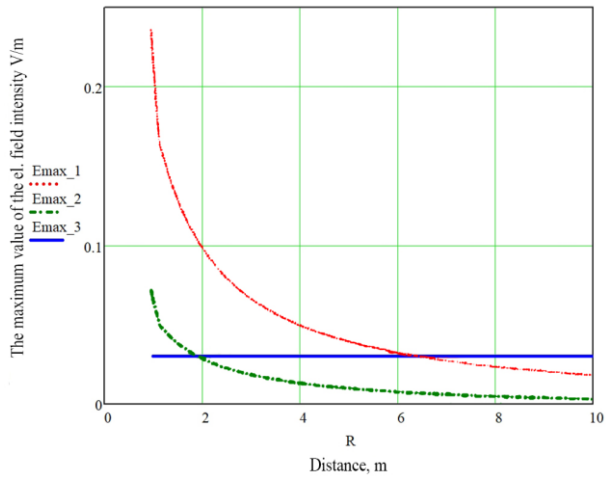
$$E_{max}(r)=\sqrt{\frac{12\pi W}{\mu\sqrt{\varepsilon\mu\omega^4}}}\cdot\left(\max\left\{\frac{1}{(2\pi\varepsilon)^2}\left(\frac{1}{r^6}+\frac{k^2}{r^4}\right)\cos^2\Theta+\frac{1}{(4\pi\varepsilon)^2}\left(\left(\frac{1}{r^3}+\frac{k^2}{r}\right)^2+\frac{k^2}{r^4}\right)\right\}\right)^{\frac{1}{2}} \tag{6}$$

In Fig. 2 shows the dependences of the maximum value of the modulus of electric field strength E_{max}, generated by relatively low-power (100 μW and 1 mW) EMR sources with frequencies of 30 and 300 MHz in a dedicated room at distances R from 1 to 10 m (100 W) sources of electromagnetic radiation, such as radio stations, 3 km away from the allocated room.

Analysis of the nature of the change in the curves shows that in the near (1 ... 8 m) zone the level of radiation from low-power sources, as expected, exceeds the level of powerful but remote sources. To realize the possibility of detecting low-power radio microphones in a complex interfering electromagnetic environment and increase the efficiency of the search system in a dedicated room are several (2 ... 4) antennas with quasi-isotropic patterns. They are installed in such a way that at any placement of the radio microphone, its distance to the antenna will be 1 ... 5 meters, which corresponds to the "near" reception area.



(a)



(b)

Fig. 2. Dependence of E max on the distance to the radiation source at the frequency: (a) 30 MHz with values of capacity of 1 mW (e_{max_1}), 100 mkW (E_{max_2}), 100 W (E_{max_3}); (b) 300 MHz with values of capacity of 1 mW (e_{max_1}), 100 mkW (E_{max_2}), 100 W (E_{max_3})

This method makes it possible to select the antenna whose output signal has the highest level. This achieves the following possibilities: to select the radiation of the RU against the background of radiation of regular radios; compensate for the non-uniformity of the radiation pattern of quasi-isotropic antennas in different spatial sectors [7].

The ability to connect to one of the inputs of the switch external ("reference") antenna significantly increases the probability of distinguishing between external and internal radiation sources in a complex electromagnetic environment, increases the speed of finding new signals.

2. Procedure for detecting sources of "unknown" electromagnetic radiation

The developed approach and the recommended procedure for detecting sources of "unknown" electromagnetic radiation, taking into account the determined and random deviations of the parameters of the electromagnetic field in the short circuit consists of the following steps:

1) The radio receiving device is tuned to a frequency range equal to the band number q of the operating range RD, q = 1, 2, ..., *Q*, $Q = \frac{RD}{\Delta F}$.
2) The antenna switch connects to the input of the device MEE "reference" antenna with the number *j* (*j = 1*).
3) Readout observed energy spectrum of the input random process based on $\hat{\sigma}^2_{clarif} = \frac{N}{N_p - N_c} \times \sum_{n \notin \Theta_c} X_n$, calculated estimation of noise intensity.
4) Narrowband signals in the frequency band are detected $\Delta F(q)$ and all components that have exceeded the threshold value are stored.
5) The antenna switch connects to the input of the device MEE antenna with a number *j = 2,..., J* after which the steps are performed in accordance with steps 3, 4.
6) For each of the detected signals $u_m(j,t), j = 2,..., J$ the number *j* and the average power for which the $\hat{P}_{j,m}$ maximum is determined, as well as $\hat{P}_{1,m}$ for the reference antenna (*j = 1*).
7) For each of the detected signals, an estimate of the difference in the observed intensity is calculated ΔP_{observ} according to the rule: $\Delta P_{observ} = 10\lg(\frac{\hat{P}_{j,m}}{\hat{P}_{1,m}}) + (\mu_c - \mu_o)$, where *j* = 2, ..., *J* with recalculation of correction factors μ_c and μ_o according to the antenna parameters in the short circuit ("signal") and "reference".
8) Determine the class of the source of electromagnetic radiation ς_m in accordance with the rule:

$$\varsigma_m = \begin{cases} \rho_{sc}, if\ \Delta P_{observ} > \Delta P_{thr2}, \\ _{notspecified}, if\ \Delta P_{thr1} < \Delta P_{observ} < \Delta P_{thr2}, \\ \rho_{ext}, if\ \Delta P_{observ} \le \Delta P_{thr1} \end{cases}$$

where ρ_{sc} – combines sources belonging to the short circuit, and ρ_{ext} – remote radiation sources.

9) Further procedure is repeated for all parts of the operating range, an action on items 1–9 for *q = 2* and so on.

The thresholds used ΔP_{thr1} and ΔP_{thr2} can be adjusted in each case taking into account the properties of the control area and the location of the receiving antennas. In addition, these thresholds may be different for different parts of the spectrum according to the actual parameters of the electromagnetic environment. Based on the above, it can be concluded that the identified structural patterns of distribution of the electromagnetic field inside the premises can be implemented to improve the use of radio devices and telecommunications to monitor the electromagnetic environment, search for devices for covert removal of confidential information and technical leakage channels.

3. Conclusions

As a result of the conducted scientific researches on the problems considered in the third section of the dissertation work the following most important scientific results are received:

1) Scientific methods of synthesis of multi-stage process of detection of technical channels of information leakage are substantiated, which includes stages: obtaining “known” EMR (I), detection of “unknown” EMR (II), identification and risk assessment of EMR and IEMRI (III) and construction of controlled perimeter zone, localization of the location of the detected source of electromagnetic radiation and counteraction to the removal (leakage) of information (IV). The results of research: "known" radiation (I), a list of "unknown" radiation (II), lists of identified EMR, parameters of digital radio signals and frequencies of detected potential leakage channels (III), coordinates of the location of the source of electromagnetic radiation in the selected room at the frequencies of the identified EMR sources and the reduction of the level of spurious emissions of the tested technical means (IV).
2) An original approach, algorithm and methods of synthesis of single – channel hardware and software search and detection of technical channels of information leakage have been developed, which solve the problem of detecting unauthorized radio microphones installed in a confined space and increase detection speed in radio interference conditions used any type of modulation.

References

[1] Atroshenko L. M. et al.: The field distribution in the near field of waveguide and horn radiators. 14 Russian Symposium with International Participation “Millimeter waves in medicine and biology” 2007, 2–5.

[2] Bakhtiiarov D. I., Konakhovych G. F., Lavrynenko O. Y.: An Approach to Modernization of the Hat and COST 231 Model for Improvement of Electromagnetic Compatibility in Premises for Navigation and Motion Control Equipment. IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), Kiev, 2018, 271–274 [http://doi.org/10.1109/MSNMC.2018.8576260].

[3] Bakhtiiarov D.: Evaluation of energy availability of means to communicate with UAVs in conditions of radioelectronic countermeasures by the enemy. Information Technology and Security 4/2016, 114–130.

[4] Gorobets N. N., Ovsyannikova Y. Y.: Wave processes in the near-field zone of large aperture antenna. 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW), Kharkiv, 2016, 1–3 [http://doi.org/10.1109/MSMW.2016.7538065].

[5] Gorobets N. N. et al.: Near-field plane distribution of rectangular waveguide excited by dominant and higher-order modes. MSMW'07 Symposium Proceedings 2007, 687–689.

[6] Kazakov G. N. et al.: Radio Monitoring of Wireless Networks Using LoRa Data Transmission Technology. 2020 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, 2020, 1–5.

[7] Kozliuk I. O. et al.: Problems of unauthorized interference to the work of uav and methods of its solving. Science-Based Technologies 30(2)/2016, 206–211.

[8] Shen X. et al.: Dynamic threshold based target signal cooperative extraction method for high frequency electromagnetic environment measurement. 3rd IEEE International Conference on Control Science and Systems Engineering (ICCSSE), Beijing, 2017, 552–555 [http://doi.org/10.1109/CCSSE.2017.8087993].

[9] Shen X. et al.: Green Hierarchical Radio-over-Fiber Distributed Antenna System Based Wireless Sensor Network for Spectrum Monitoring. IEEE International Conference on Computer and Information Technology, Xi'an, 2014, 880–883 [http://doi.org/10.1109/CIT.2014.105].

[10] Vedenkin D. A., Potapova O. V., Sedelnikov Yu. Ye.: Antennas focused in the near radiated field zone. Features and technical application. International Conference on Antenna Theory and Techniques 2013, 560–565.

[11] Yurtsev O. A., Naumovich N. M.: Field research in near zone of antennas carried out at Belarusian state University of informatics and radioelectronics. International Conference on Antenna Theory and Techniques 2013, 566–571.

[12] Zhong L.: Monitoring Function Design of Radio Monitoring Management System Based on C/S Architecture. 10th International Conference on Information Technology in Medicine and Education (ITME), Qingdao, China, 2019, 424–428 [http://doi.org/10.1109/ITME.2019.00101].

M.Sc. Denys Bakhtiiarov
e-mail: bakhtiiaroff@tks.nau.edu.ua

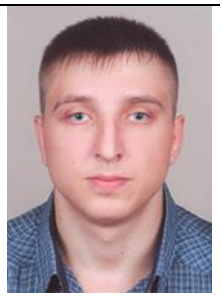
Position: graduate student of National Aviation University. Research interests: information security, information coding. Publications: more than 30 scientific publications.

http://orcid.org/0000-0003-3298-4641

M.Sc. Oleksandr Lavrynenko
e-mail: oleksandrlavrynenko@tks.nau.edu.ua

Position: graduate student of National Aviation University. Research interests: signal processing, speech recognition. Publications: more than 30 scientific publications.

http://orcid.org/0000-0002-3285-7565



M.Sc. **Nataliia Lishchynovska**
e-mail: natashalil858@ukr.net

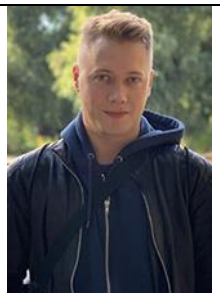
Position: graduate assistant of the Department of Information Security Means. Research interests: method for the synthesis of high-quality resonators. Publications: 13 scientific publications.

http://orcid.org/0000-0002-1913-8419

M.Sc. Ivan Basiuk
e-mail: basya2000@gmail.com

Position: Senior Devops Engineer, Globallogic LTd., Kyiv, UA
Publications: more than 20 scientific publications.

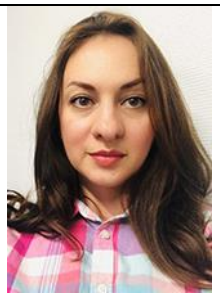
http://orcid.org/0000-0002-2666-1136



M.Sc. Tetiana Prykhodko
e-mail: tata@mirohost.net

Position: Head of Support Department, Internet Invest, Ltd., Kyiv, UA
Publications: more than 20 scientific publications.

http://orcid.org/0000-0001-6909-7697



otrzymano/received:14.12.2020 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2581

METHOD FOR DETERMINING THE ACTUAL PRESSURE VALUE IN A MV VACUUM INTERRUPTER

Michał Lech, Damian Kostyła
Lublin University of Technology, Faculty of Electrical Engineering and Computer Science, Lublin, Poland

Abstract. *The paper describes the author's method for determining the actual pressure value in a vacuum interrupter of a MV disconnector during high-voltage laboratory tests. The need to develop such a method was due to the lack of possibility to measure the pressure directly in the vacuum interrupter during the tests. There was a risk of damaging the vacuum gauge as a result of an electrical jump. The proposed method consists in measuring, in de-energized conditions, the difference in pressure between a set of vacuum pumps and a prototype vacuum interrupter made for the purpose of implementing this method. Thus, knowing the pressure value at the pumps and having the scaling characteristics of the system determined, it will be possible to determine the actual pressure inside the tested MV disconnecting extinguishing interrupter during high-voltage tests. Measurements of pressure drop in the pumping channel were carried out for air and three electro-negative gases: helium, argon and neon, used in the study by the authors of this paper.*

Keywords: vacuum switchgears, vacuum technology, vacuum systems, pressure measurement

METODA OKREŚLANIA RZECZYWISTEJ WARTOŚCI CIŚNIENIA W PRÓŻNIOWEJ KOMORZE GASZENIOWEJ SN

Streszczenie. *W artykule opisano autorski sposób pozwalający na określenie rzeczywistej wartości ciśnienia w próżniowej komorze rozłącznikowej SN podczas wysokonapięciowych badań laboratoryjnych. Konieczność opracowania metody tego typu wynikła z braku możliwości pomiaru ciśnienia bezpośrednio w komorze próżniowej w trakcie badań. Istniało bowiem ryzyko uszkodzenia próżniomierza w wyniku przeskoku elektrycznego. Proponowana metoda polega na pomiarze w warunkach beznapięciowych różnicy ciśnienia pomiędzy zestawem pomp próżniowych, a prototypem komory próżniowej wykonanym dla potrzeb realizacji tej metody. Dzięki temu znając wartość ciśnienia przy pompach oraz mając wyznaczone charakterystyki skalowania układu, w trakcie badań wysokonapięciowych możliwe będzie określenie rzeczywistego ciśnienia wewnątrz badanej rozłącznikowej komory gaszeniowej SN. Przeprowadzono pomiary spadku ciśnienia w kanale pompowym dla powietrza oraz trzech gazów elektroujemnych: helu, argonu oraz neonu, wykorzystywanych w badaniach przez autorów niniejszego artykułu.*

Słowa kluczowe: próżniowa aparatura łączeniowa, technologia próżniowa, systemy próżniowe, pomiar ciśnienia

Introduction

The continuous increase in demand for electricity determines the development of medium voltage lines, both in terms of their length and technological advancement. Currently, there are over 306 thousand kilometers of medium voltage lines in Poland [6]. For efficient and failure-free operation of power infrastructure, appropriate switching devices are necessary, maintenance-free operation, thus ensuring improved reliability indices, the values of which are billed to Distribution System Operators (DSOs) [1, 5, 7].

Switching devices used as components of medium voltage lines are divided into two types. The first are open-type devices and the continuously gaining in importance closed-type devices. The most common closed devices are SF6 gas and vacuum arc extinguishers. Sulfur hexafluoride is an extremely harmful greenhouse gas, as its GWP100 is 22,000 times greater than that of carbon dioxide. In order to limit the use of this medium, the Kyoto Protocol was signed in December 1997, which obliged the countries of the world to reduce the amount of greenhouse gases produced [2, 8].

An alternative to SF6 gas is apparatus based on vacuum technology. Over the last few years, there has been a definite increase in the number of devices installed based on vacuum technology. This is due to the excellent performance of vacuum for arc extinguishing and its almost neutral environmental impact. Figure 1 shows the percentage of newly designed circuit breakers in medium voltage networks installed in the world power industry over the last decades.

The aforementioned factors have made the switchgear based on vacuum technology become the main development trend dedicated to medium voltage lines, which is perfectly exemplified by the innovative vacuum disconnector EKTOS, developed by a team of scientists from the Lublin University of Technology in cooperation with the EKTO company from Białystok, dedicated to intelligent medium voltage networks of the Smart Grid type [3, 10, 11].

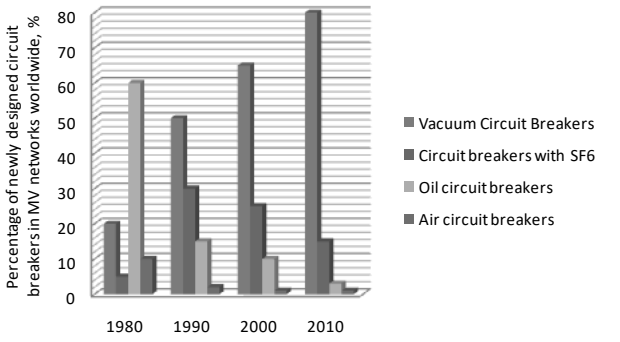


Fig. 1. Percentage share of MV circuit breakers in electric power network in the world between 1980 and 2010 (own elaboration based on [9])

Vacuum switching devices are based on vacuum interrupters used for connecting current circuits of a given device. The construction of MV vacuum interrupter consists of ceramic or glass casing, in which the current track (supply and contact pair) is located. The material and shape of the contacts is selected at the design stage in such a way as to limit the wear and tear of contact pads during switching operations. The possibility of movement of the mobile contact results from the use of a bellows, which ensures the tightness of the system. Inside the interrupter there is also a condensation screen on which conductive particles from the arc discharge occurring between the contacts are deposited. This element is necessary because conductive particles deposited on the interrupter casing could reduce the dielectric strength. A similar shield is located at the spring bellows. The manufacture of vacuum interrupters is a high-tech process and requires the use of special materials to withstand differential pressures and mechanical shocks. The pressure inside SN vacuum extinguishing interrupters oscillates around 10^{-4} Pa.

Figure 2 shows the structure of a typical MV vacuum interrupter used in power switchgear, while Table 1 presents the basic technical parameters of vacuum interrupters from the largest manufacturers.

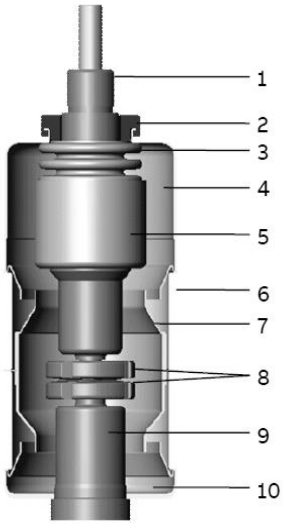


Fig. 2. Construction of the vacuum interrupter (1 – stem / terminal, 2 – twist protection, 3 – metal bellows, 4 – interrupter lid, 5 – shield, 6 – ceramic insulator, 7 – shield, 8 – contacts, 9 – stem / terminal, 10 – interrupter lid)

Table 1. Basic technical parameters of MV vacuum interrupters of the largest manufacturers

Manufacturer	Rated voltage, kV	Rated continuous current	Mechanical durability, thou.
ABB	7.2 – 36	400 ÷ 3150	10 ÷ 1000
EATON	12 – 17.5	630 ÷ 3200	10
SIEMENS	7.2 – 36	630 ÷ 4000	30
ITR (Poland)	≤ 24	400	2

1. Motivation to conduct the study

The growing popularity of vacuum as an insulating medium in power equipment and the need to improve its operating parameters has prompted the authors of this paper to undertake research aimed at developing a method to increase the switching capacity and electrical strength of quenching interrupters dedicated to modern switching equipment. The planned research concerns mainly a comparative analysis of changes in dielectric strength of the contact system of a vacuum interrupter in the open state, depending on the contact material used as well as on the type and pressure of gas in the vacuum interrupter. This pressure is the key physical quantity occurring during the implementation of tests in the Department's Switching and Switchgear Laboratory. It is therefore important to correctly determine this quantity, on which a number of technical parameters depend.

Unfortunately, because testing the dielectric strength of interconnects requires the use of high test voltages, it is not possible to measure the pressure directly in the vacuum interrupter under test. Attempting such measurements would create the possibility of damaging the test head through the possibility of an electrical jump to the vacuum gauge element. Therefore, the authors of this paper developed a method to determine the pressure inside the SN vacuum interrupter under test from the pressure measured at the vacuum pumps. The volume of the whole article should include an even number of pages. The last page should be filled at least 50%. The author should make

2. Test stand

A special version of the vacuum interrupter was designed and constructed for the purpose of this study. It reproduces the geometrical parameters of a real medium-voltage vacuum interrupters used in MV switching devices (Fig. 3). The interrupter is equipped with a vacuum connection ferrule at the height of its contacts, thanks to which it is possible to connect a measuring head there. Obtaining the specified pressure value inside the interrupter was ensured by its adaptation to the connection of the pumping channel.

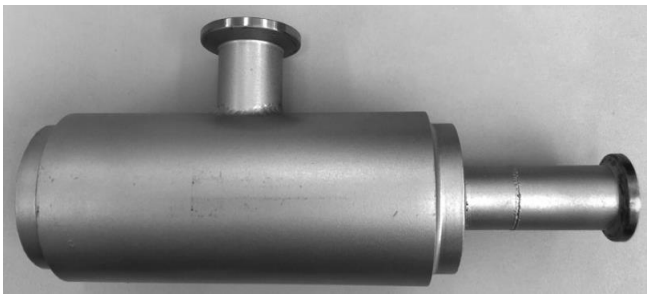


Fig. 3. Vacuum interrupter with special design

A schematic diagram of the test stand developed for the determination of the actual pressure value in the SN disconnecting vacuum interrupter is shown in Figure 4, while a view of the stand is shown in Figure 5.

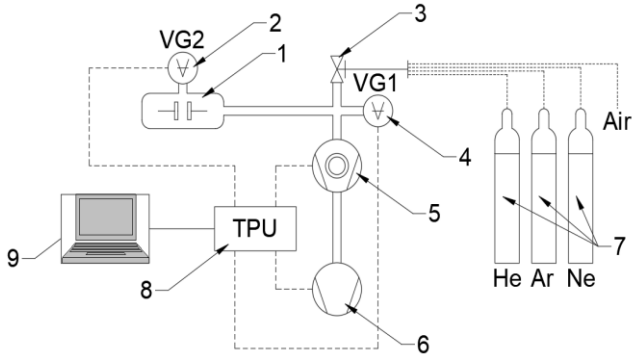


Fig. 4. Scheme of test stand for pressure measurement (1 – prototype of vacuum interrupter in special design, 2 – vacuum gauge no. 2, 3 – vacuum hand valve, 4 – vacuum gauge no. 1, 5 – turbomolecular pump, 6 – rotary pre-pump, 7 – set of technical gases, 8 – TPU control unit, 9 – PC station)

Fig. 5. View of the laboratory station for pressure measurements

The developed test stand consists primarily of a set of vacuum pumps: a rotary pre-pump and a turbomolecular pump. This set operates at a capacity of 90 l/s, thanks to which it is possible to obtain vacuum in the interrupter at the value of the pressure used in currently manufactured SN quenching interrupters. Special vacuum valves used for dosing technical gases or for aeration of the system cooperate with this system. Pressure measurement in the system is carried out using vacuum measuring heads installed at the vacuum pump set (VG1) and at the vacuum interrupter (VG2). The vacuum pumps and measuring heads are operated by a TPU control module connected to the laboratory computer network. The control and reading of the pressure values from the vacuum gauges is done using a computer unit and dedicated software provided by the manufacturer of the vacuum set.

3. Measurement results

The idea of the research conducted was to determine the pressure drop between a vacuum pump and the subject vacuum interrupter prototype.

Measuring the pressure of gases other than air using measuring heads requires correcting the read value by applying calibration factors. Below a certain pressure value, the measuring head operates in the ionization (penning) head mode, where the value read is equal to the actual pressure value. Above this value, the measuring head switches to thermal head mode (Pirani). In this mode, the actual pressure value is linearly dependent on the actual pressure value indicated by the vacuum meter. This is determined by the following relation:

$$p_{eff} = C \times p_{reading} \tag{1}$$

where p_{eff} is the actual pressure value, C denotes the calibration factor and $p_{reading}$ the pressure value read.

The calibration factors for the industrial gases used are shown in Table 2.

Table 2. Calibration factors for test gases [4]

Gas type	Calibration factor	Valid range
He	1.40	$3\times10^{-1} \div 3\times10^{1}$ Pa
Ar	1.57	$3\times10^{-1} \div 1\times10^{2}$ Pa
Ne	1.70	$3\times10^{-1} \div 3\times10^{1}$ Pa
Air	1.00	$3\times10^{-1} \div 3\times10^{1}$ Pa

Above the pressure value specified by the range in Table 2, the actual pressure value should be determined using the characteristics provided by the measuring head manufacturer. However, this pressure range is not of interest to the authors of this paper.

In order to determine the pressure drop across the pumping channel between the vacuum pump set-up and the prototype vacuum interrupter, pressure values were measured at two locations in the system, using vacuum gauges VG1 and VG2. Measurements were made for air and three noble gases: helium, argon and neon, in the pressure range 10^{-4} Pa ÷ 10^{1} Pa.

Based on the obtained measurements, the characteristics $p_{VG2} = f(p_{VG1})$ were developed and are shown in Figures 6–9.

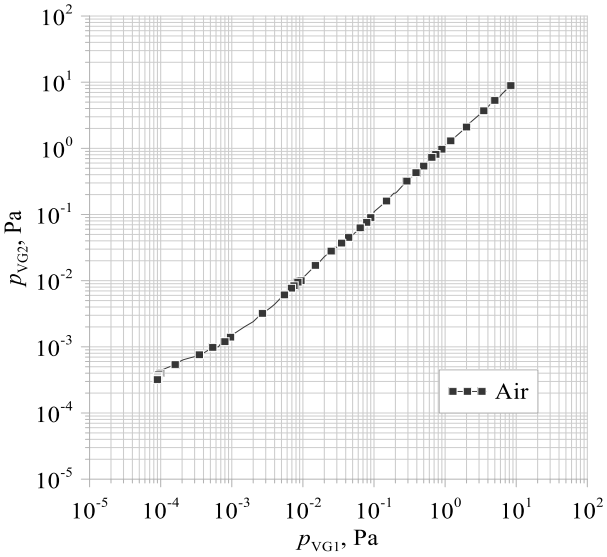


Fig. 6. Characteristics of pressure measured by vacuum gauge VG1 as a function of pressure measured by vacuum gauge VG2 for air

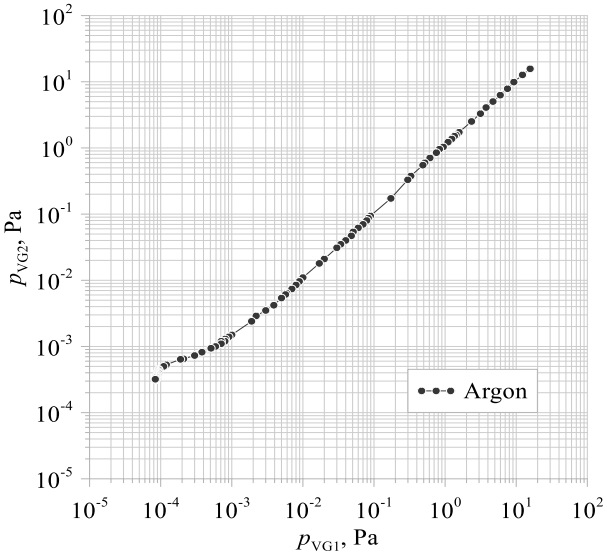


Fig. 7. Characteristics of pressure measured by vacuum gauge VG1 as a function of pressure measured by vacuum gauge VG2 for argon

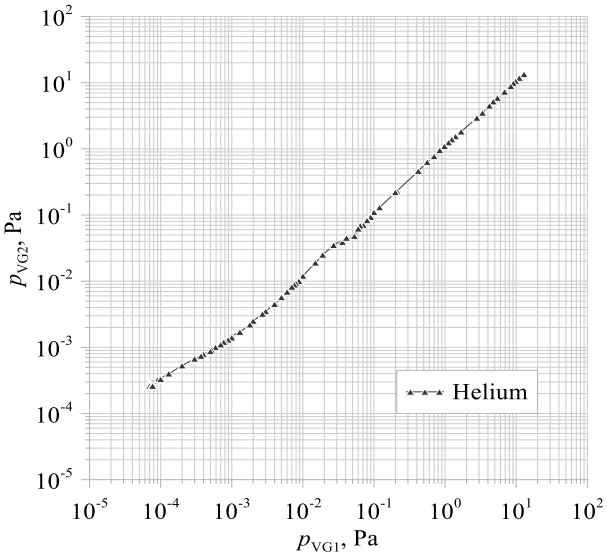


Fig. 8. Characteristics of pressure measured by vacuum gauge VG1 as a function of pressure measured by vacuum gauge VG2 for helium

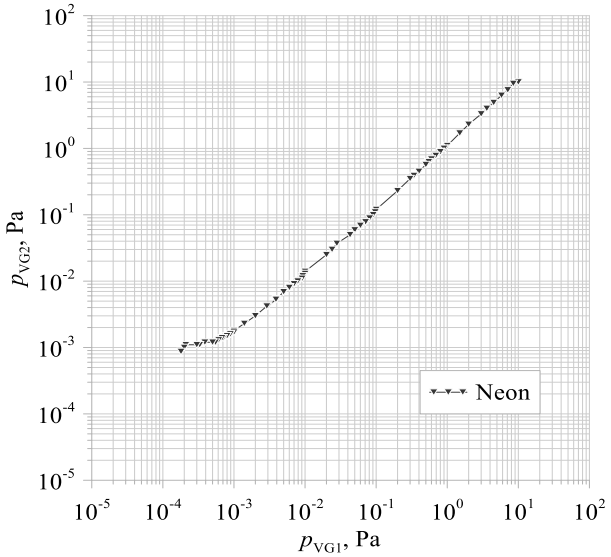


Fig. 9. Characteristics of pressure measured by vacuum gauge VG1 as a function of pressure measured by vacuum gauge VG2 for neon

Analyzing the obtained characteristics, it can be observed that above a pressure value of about 10^{-2} Pa, the pressures indicated by vacuum gauges VG1 and VG2 are approximately equal. Below this value, the pressures indicated by the measuring head VG2 are higher compared to VG1. This is due to the difficulty of maintaining high vacuum along the length of the pump channel, where gaskets and channel diameter changes occur. These locations are highly sensitive and present potential opportunities for air ingress into the system. Therefore, in the pressure range of 10^{-4} Pa ÷ 10^{-2} Pa, the pressure difference between vacuum gauges VG1 and VG2 was determined, which will be used to determine the actual pressure in the tested disconnecting interrupter in the authors' further high-voltage research.

4. Conclusions

The dynamic development seen in the power industry is associated with the need to develop new devices with better performance, while taking into account the trends of miniaturization. This entails a number of works carried out by research teams around the world.

The research team working at the Faculty Laboratory of Switchgear and Distribution Equipment is currently working on widely understood improvement of technical parameters of vacuum interrupters dedicated to modern switchgear used in smart grids.

In the course of high-voltage research work, the necessity arose to determine the actual pressure prevailing in the tested vacuum interrupter in a manner that does not expose the vacuum meter to damage associated with direct pressure measurement near the contact system. A measurement method was developed based on the use of two measuring heads and a prototype of a vacuum SN extinguishing interrupter. Based on the measurement of the pressure values at two locations in the system, the difference in the readings of the vacuum gauges was determined, which will be used to determine the residual gas pressure in further high-voltage tests of the disconnecting interrupters based on the pressure prevailing at the vacuum pump set.

References

[1] Konarski M., Węgierek P.: The use of power restoration systems for automation of medium voltage distribution grid. Przegląd Elektrotechniczny 7/2018, 167–172.
[2] Kyoto Protocol to the UN Framework Convention on Climate Change. Kyoto 1997.
[3] Lech M.: Disassemblable vacuum interrupter as an innovative test stand designed for research on improving the operational parameters of power switching apparatus. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOS 3/2020, 66–69.
[4] Leybold company data sheets.
[5] Ordinance of the Minister of Economy of 4 May 2007 on detailed conditions of the power system operations.
[6] Power engineering, distribution and transmission, Polish Power Transmission and Distribution Association's Report, Poznań, 2020.
[7] Quality Regulation 2018 – 2025 for Distribution System Operators.
[8] Regulation (EU) No 517/2014 of the European Parliament and of the Council of 16 April 2014 on fluorinated greenhouse gases.
[9] Slade P. G.: The Vacuum Interrupter Theory, Design, and Application, CRC Press 2007.
[10] Węgierek P., Lech M.: Test stand for testing and diagnostics of medium voltage vacuum interrupters. Przegląd Elektrotechniczny 2/2021, 176–179.
[11] Węgierek P., Staszak S., Pastuszak J.: EKTOS – innovative medium voltage outdoor vacuum disconnector in a closed housing dedicated to the Network smart grids. Wiadomości Elektrotechniczne 11/2019, 21–25.

M.Sc. Eng. Michał Lech
e-mail: m.lech@pollub.pl

Michał Lech (born 1995, Kielce) – Ph.D. student at the Doctoral School at the Lublin University of Technology. His scientific interests include innovative methods of designing and constructing devices operating in Smart Grid networks which increase the reliability of electricity transmission and distribution. His current scientific research concerns improving the operational parameters of SF6 free switching equipment.

http://orcid.org/0000-0002-4732-2459

M.Sc. Eng. Damian Kostyła
e-mail: kostyla.damian@gmail.com

Damian Kostyła (born 1995, Świdnik) – A graduate of electrical engineering at Lublin University of Technology. His research interests include modern methods of designing power equipment for medium voltage Smart Grids.

The current research is on increasing the dielectric strength of contact gaps in MV vacuum interrupters.

http://orcid.org/0000-0002-9012-0158

otrzymano/received: 28.02.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2455

OVERVIEW OF FEATURE SELECTION METHODS USED IN MALIGNANT MELANOMA DIAGNOSTICS

Magdalena Michalska
Lublin University of Technology, Department of Electronics and Information Technology, Lublin, Poland

Abstract. *Currently, a large number of trait selection methods are used. They are becoming more and more of interest among researchers. Some of the methods are of course used more frequently. The article describes the basics of selection-based algorithms. FS methods fall into three categories: filter wrappers, embedded methods. Particular attention was paid to finding examples of applications of the described methods in the diagnosis of skin melanoma.*

Keywords: feature selection methods, filter methods, wrappers methods, embedded methods

PRZEGLĄD METOD SELEKCJI CECH UŻYWANYCH W DIAGNOSTYCE CZERNIAKA

Streszczenie. *Obecnie stosuje się wiele metod selekcji cech. Cieszą się coraz większym zainteresowaniem badaczy. Oczywiście niektóre metody są stosowane częściej. W artykule zostały opisane podstawy działania algorytmów opartych na selekcji. Metody selekcji cech należące dzielą się na trzy kategorie: metody filtrowe, metody opakowujące, metody wbudowane. Zwrócono szczególnie uwagę na znalezienie przykładów zastosowań opisanych metod w diagnostyce czerniaka skóry.*

Słowa kluczowe: metody selekcji cech, metody filtrowania, metody opakowujące, wbudowane metody

Introduction

Early detection and classification of melanoma is extremely important for treatment and patient outcome. In order to classify selected features of the image, they must be properly selected. Important in the diagnostic processes is the selection of an appropriate set of data (dermatoscopic images), a classification method of skin lesions, the classification process and selection of features. This last stage is also not the easiest one. Figure 1 presents a diagram of the diagnostic process based, of course, on an appropriately selected method of selecting features.

There are many methods of selecting features. The feature selection methods are broken down into three basic categories: filters, wrappers and embedded methods [4]. In recent years, researchers have developed many methods to select features through IT tools [6, 11, 32]. Still new feature selection methods are being proposed.

The rapidly increasing number of features is a very serious problem to be solved. This increases the computational complexity of the algorithm, extends the learning process and increases multi-level classification method.

The best result of the classifier is given by a properly selected feature selection algorithm. Feature selection, reduction of the feature space dimensionality, reduces the number of free parameters in the classifier necessary for estimation. When collecting data again, you can focus only on the features important for the classification algorithm [31]. Filters mainly use the general characteristics of data sets. Wrappers and embedded methods build a subset of functions based on selected algorithms.

The most important algorithms for selecting the features of medical images include methods [21]: SBS (Sequential Backward Selection), SFS (Sequential Forward Selection) and its modifications (SFFS (Sequential Forward Floating Search)). The SFFS algorithm requires providing the algorithm's stop condition, the number of necessary operations does not have to be so large due to the removal of features previously selected from the subset.

Other methods are: method Plus-L-Minus-R, NNFP (Nearest Neighbor with Feature Projection), methods based on genetic algorithms, OSA (Oscillating Search Algorithm), methods based on the use of fractal dimension, methods based on information thery.

Figure 2 presents a summary of used groups of feature selection methods based on four categories such as classification, segmentation, annotation and retreval. Scientists use filter methods the most, followed by embeded methods. Filter methods are at the first place of use.

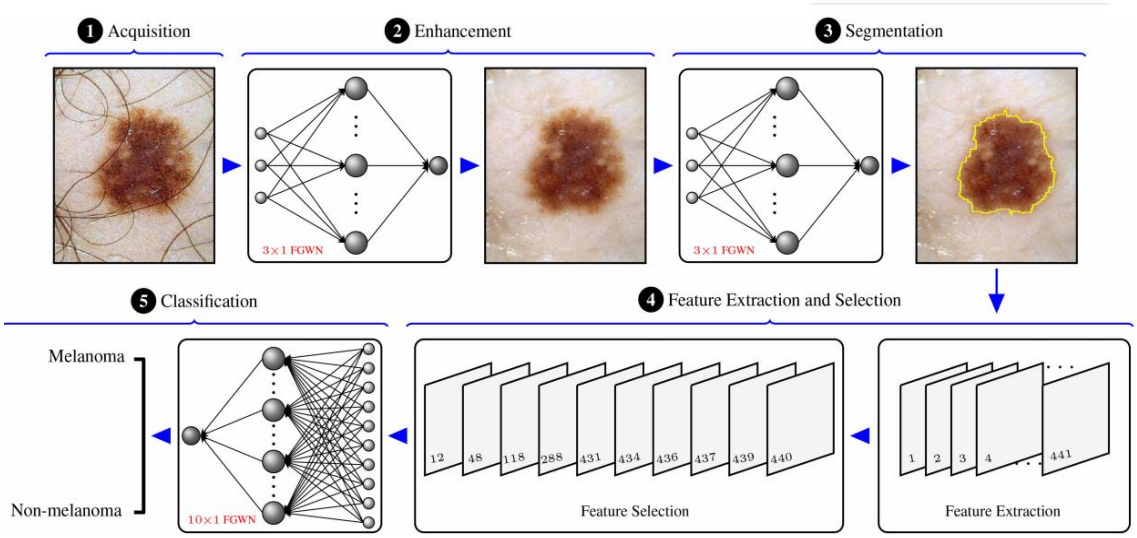


Fig. 1. Scheme of diagnosis method of skin lesions from dermoscopic images [29]

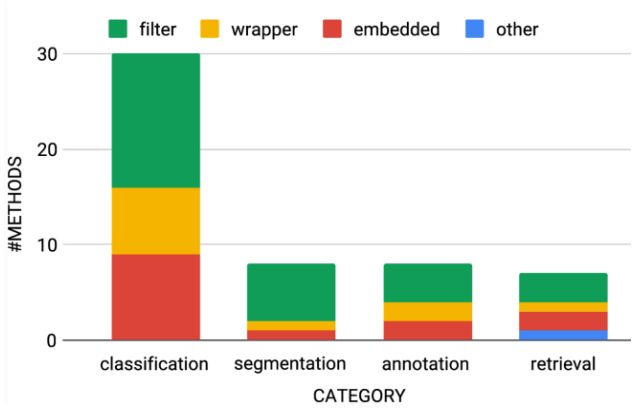


Fig. 2. Application of selected types of feature selection methods in numbers [4]

1. Filters methods

As a pre-processing process, the most frequently used by researchers are filter methods. The methods use a statistical measure and the functions are selected for retention or removal from the data. The methods are usually one-dimensional and take into account features independently or in relation to the dependent variable. The filters methods include: correlation-based (CFS) [13], consistency-based filter [7] and information gain [12, 13]. CFS then combines this evaluation formula with an appropriate measure of correlation and a heuristic filter search strategy, selecting subsets of attributes not correlated between them. It can showe a correlation with the all class.

Another very common filters method is ReliefF [18, 19]. Filter models according to developed sources are more computationally efficient [39]. The relief algorithm is effective in determining a given feature [20]. Figure 3 gives a detailed description of the reliefF algorithm. ReliefF randomly choses an instance R_i from class. It can find K for the nearest neighbors from the same class (nearest hits H) and from the different classes (nearest misses M), i.e.:

$$W_i = W_i - \frac{\sum_{k-1}^{K} D_H(k)}{n \cdot k} + \sum_{c=1}^{C-1} p_c \cdot \frac{\sum_{k=1}^{K} D_M(k)}{n \cdot k} \quad (1)$$

where W_i – quality measure for feature according to R_i values, hits H and misses M; $D_H(k)$ and $D_M(k)$ – distance between the selected instance and its nearest neighbors in H (or M); c_p – class probability c; n – repeats n times.

Input: Feature data matrix: D, repeat times: n, the number of the neighbors: K
Output: Vector W for the feature attributes ranking
Begin
  for j=1 to n do
    Randomly select an instance R_j;
    Find K nearest hits H and nearest misses M;
    for i=1 to all features do
      Updating estimation W_i by Equation(1);
    end
  end
End

Fig. 3. ReliefF algorithm [39]

In the study [23], several different feature selection algorithms were used to create subsets for classifiers. The algorithms are based on various bases, e.g. Pearson's correlation coefficient based on feature selection gain factor [34]. Relief-F, principal component analysis (PCA) and feature selection based on correlation (CFS) are also used in many works. These algorithms are commonly used, because they have a number of advantages. Computing performance is one of them. Additionally, they have become less time-consuming and do not result in excessive and independent evaluation criteria [28].

In most cases, the selected features are determined by the correlation results of statistical tests [16]. Common used defining correlation coefficients are [1]: pearson's correlation, LDA (Linear Discriminant Analysis), ANOVA stands, Chi-Square [29]. The data based on the dermoscopic images served as a test kit to evaluate the effectiveness of the classification.

Several feature selection algorithms were used in [29]: ReliefF algorithm, Fisher score [5], chi-square. Table 1 below shows a comparison of the results from [29].

Table 1. Accuracy for different feature selector [29]

Parameter	Feature selector				
	ReliefF	FCBF	FS	mRMR	Chi-squere
Mean of accuracy [%]	87.1	85.8	85.8	87	85.8

In filtering methods, class separation, error probability, and inter-class distance are used. Very common is correlation-based feature selection, entropy, consistency-based feature selection and filter methods do not remove multicollinearity, it should be fixed before training models [17].

2. Wrappers methods

The scheme of functioning has been re-colored in Figure 4 wrappers methods. It belongs to them set of all features. Next is selected the best subset to generate a subset and learnig algorithm is started. After all those the Performance is been done.

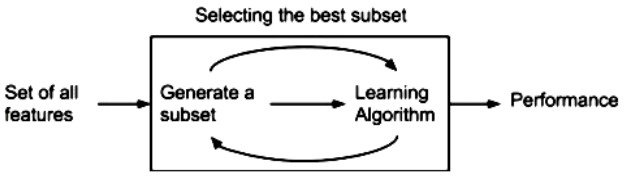


Fig. 4. Wrapper methods model [14]

In wrapper methods important is to use a subset of features and train a model. Based on the inferences from the previous model, features subset are added or removed [5, 9]. Very common for wrapper methods are forward feature selection, backward feature elimination, recursive feature elimination. These are usually computationally very expensive.

The study [24] adopted them. Greedy stepper search methods contain subsets in forward or backward direction. The selection stop when any feature is added or removed. This function degrades the result of the subset up to this point [37]. The best first method searches for subsets of functions. An empty feature set starts the selection forward, features compatible with the evaluation method are added to the data set. On the other hand, all features start backselection, and mismatched features are removed from the set [15].

The Support Vector Machine (SVM) recursive support vector machine-recursive feature elimination (SVM-RFE) method, which is a very typical wrapper selector. The method was first developed for the gene selection process using the SVM classifier [20, 25, 26]. The system of mobile applications [8] helps to classify skin nevi on dermatoscopic images as melanoma, benign nevi.

A type of machine learning technique is Genetic Programming (GP). It allows the use of the evolutionary algorithm for simple and understandable classifiers [22, 36]. GP is also used to diagnose tumor expression, c was used in the selection of features and classifiers [2, 37].

In order to use the selected skin lesion classification algorithm, first of all, reduce the size of the dermatological image on which it is located. With a large amount of data, it is useful to reduce features and design functions to reduce their size. This operation allows for greater efficiency of the used classifier. In the work [33], an innovative, two-step GP algorithm

was developed to select the features and structure of features for the classification of the skin cancer picture. The local binary pattern helps to show gray and color characteristics from dermoscopic images.

Unlike wrapper and embedded methods, filter methods require more computational effort. In addition, they are less accurate in their selection. Wrappers are over-matched when the number of samples is smaller than the number of elements.

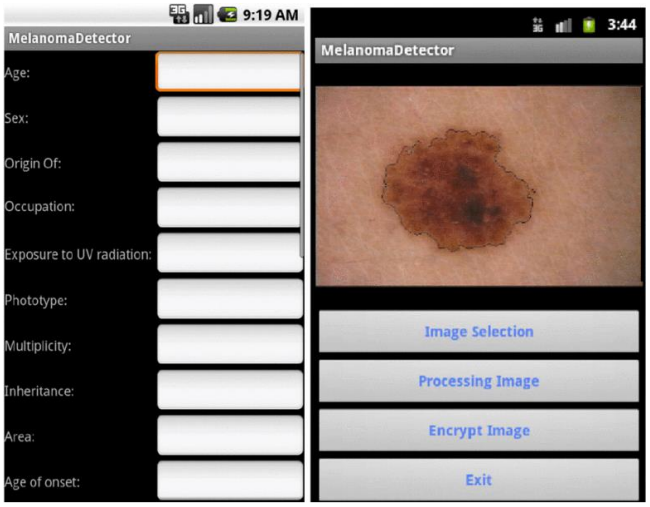


Fig. 5. Mobile application from melanoma detection [8]

3. Embedded methods

Embedded methods use internal representations of selected classifiers, which evaluate the usefulness of features in the learning process. In order to build a model, methods of selecting features are also used. [36]. They also usually give better results than filter methods. They are designed according to the selected classification algorithm. The methods are faster because the selection process does not require calling the classifier multiple times for each feature subset.

For detection melanoma the best is to find combination of different criteria. The lesion area was analyzed in terms of lesion area division parameters – Figure 6 [38].

Benign lesions differ from malignant ones in terms of selected characteristic attributes. The analyzed homogeneity and selected color characteristics (Figure 3a) usually have higher values in the case of benign skin lesions. in order to classify with the highest efficiency, a combination of selected attributes should be used [9].

One of the rapidly developing methods is LASSO (Least Absolute Shrinkage and Selection Operator) [3] and method RLS (Relaxed Linear Separability) [10]. The methods are especially used with bigger number of samples in the training set. The following are used to reduce the dimensions of the function: Sammon mapping, principal components analysis (PCA), decomposition of singular values. Very often a low variance filter and a high correlation filter are used. Useful for feature selection in addition to classifiers are Random forests [30, 35].

The diagnosis of melanoma is possible thanks to the visualization of the analysis of structured data [30]. The created data set made to measure significantly exceeds the limits of today's multi-dimensional and multi-dimensional visualization techniques. Visualization based on (PCA) [16] reduces dimensionality to a manageable range and provides better visualization. PCA may introduce errors, but the tolerance of error can be assessed and controlled [27].

In [1] using SVM based on the selected features from PCA, achieved accuracy of around 92% with 11 features. Figure 7 shows the selection results for 5 features using the PCA method. the developed methods allow to distinguish malignant from benign changes, becoming a fairly powerful diagnostic tool.

Sequential forward search algorithms SFS (ang. sequential forward selection) and sequential search backwards SBS (ang. Sequential Backward Selection) are examples of simple boxing methods. In the case of the first method, the algorithm adding a new feature in each subsequent step [2]. With both of these methods, the forward or reverse scanning step is followed by a reverse scanning step. This allows for the removal of a feature in the SFS algorithm that becomes redundant after adding others, and in the case of the SBS algorithm. It is possible to consider a given feature again, although it was removed in an earlier step of the algorithm [32]. Also very common is Backward Feature Elimination and inverse process Forward Feature Construction [35].

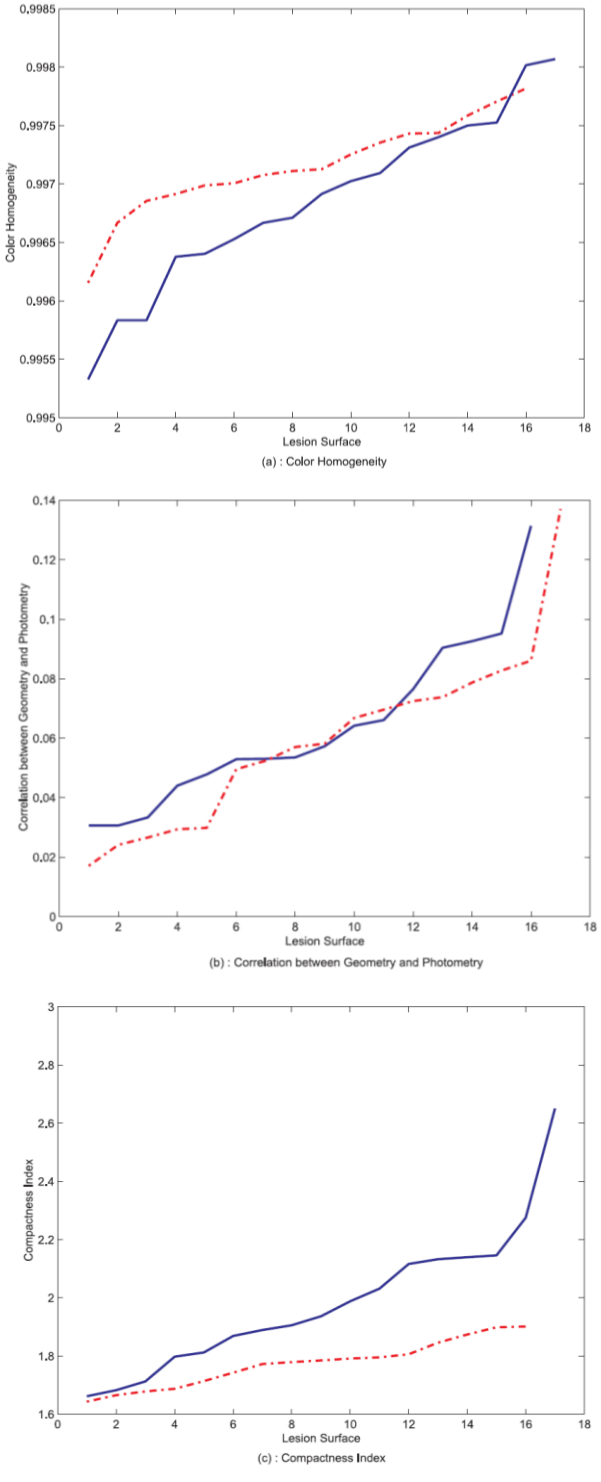


Fig. 6. Correlations and compactness in relation to benign lesions "red dotted line" and malignant lesions "blue solid line" [38]

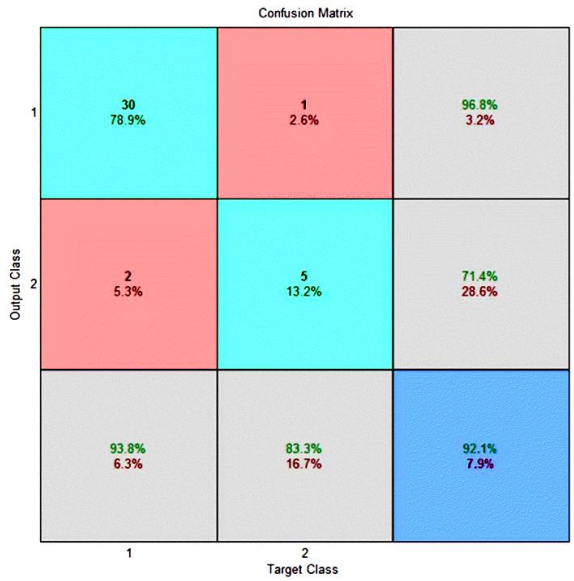


Fig. 7. The matrix of using SUV for features selction using PCA [1]

4. Conclusions

Function selection algorithms use many features of skin lesions. There are often a number of characteristics that need to be kept. The feature subset evaluator measures a features quantity and returns the search value. The choice of features and their design alone. Performance improvements can be achieved by selecting an appropriately selected feature, or by using a more extensive multi-level feature. Each of the methods mentioned have many advantages and disadvantages. Their features complement each other. Filter methods are less computationally expensive, embedded methods allow for a more precise selection. The use of additional neuron networks, SUVs, methods of feature elemination, decision trees gives the opportunity to obtain the best possible result, allowing for an accurate diagnosis.

References

[1] Alquran H., Qasmieh I. A., Alqudah A. M., Alhammouri S., Alawneh E., Abughazaleh A., Hasayen F.: The melanoma skin cancer detection and classification using support vector machine. IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017, 1–5 [http://doi.org/10.1109/AEECT.2017.8257738].

[2] Al-Sahaf H., Al-Sahaf A., Xue B., Johnston M., Zhang M.: Automatically Evolving Rotation-Invariant Texture Image Descriptors by Genetic Programming. IEEE Transactions on Evolutionary Computation 21(1)/2017, 83–101.

[3] Andersen S. W., Runger G. C.: Automated feature extraction from profiles with application to a batch fermentation process. Journal of the Royal Statistical Society: Series C (Applied Statistics) 61(2)/2012, 327–344.

[4] Bolón-Canedo V., Remeseiro B.: Feature selection in image analysis: a survey. Artif Intell Rev 53/2020, 2905–2931.

[5] Celebi M. E., Aslandogan Y. A., Stoecker W. V., Iyatomi H., Oka H., Chen X.: Unsupervised border detection in dermoscopy images. Skin Res Technol. 13/2007, 1–9.

[6] Chmielnicki W.: Efektywne metody selekcji cech i rozwiązywania problemu wieloklasowego w nadzorowanej klasyfikacji danych. Rozprawa doktorska. Instytut Podstawowych Problemów Techniki PAN, Kraków 2012.

[7] Dash M., Liu H.: Consistency-based search in feature selection. Artificial Intelligence 151(1–2)/2003, 155–176 [http://doi.org/10.1016/S0004-3702(03)00079-1].

[8] Doukas C., Stagkopoulos P., Kiranoudis C. T., Maglogiannis I.: Automated skin lesion assessment using mobile technologies and cloud platforms. Engineering in Medicine and Biology Society (EMBC) – Annual International Conference of the IEEE, 2012.

[9] Ercal F., Chawla A., Stoecker W.V., Lee H., Moss R. H.: Neural Network diagnosis of malignant melanoma from color images. IEEE Transactions on Biomedical Engineering 41(9)/1994, 837–845.

[10] Gościk, J., Łukaszuk, T.: Application of the recursive feature elimination and the relaxed linear separability feature selection algorithms to gene expression data analysis. Advances in Computer Science Research 10/2013, 39–52.

[11] Guyon I., Elisseeff A.: An introduction to variable and feature selection. Journal of Machine Learning Research 3/2003, 1157–1182.

[12] Hall M., Smith Lloyd A.: Practical feature subset selection for machine learning. Springer 1998.

[13] Hall M.: Correlation-based feature selection for machine learning. Department of Computer Science 19/2000.

[14] https://moredvikas.wordpress.com/2018/10/09/machine-learning-introduction-to-feature-selection-variable-selection-or-attribute-selection-or-dimensionality-reduction/

[15] Huang J., Ling C. X.: Using AUC and accuracy in evaluating learning algorithms. IEEE Trans Knowledge Data Eng. 17(3)/2005, 299–310.

[16] Keerthi Vasan K., Surendiran B.,: Dimensionality reduction using Principal Component Analysis for network intrusion detection. Perspectives in Science 8/2016, 510–512.

[17] Khan M. A., Tallha A., Muhammad S., Aamir S., Khursheed A., Musaed A., Syed I. H., Abdualziz A.: An implementation of normal distribution based segmentation and entropy-controlled features selection for skin lesion detection and classification. BMC Cancer 18(1)/2018, 638.

[18] Kira K., Rendell L. A.: A practical approach to feature selection. Machine Learning Proceedings 1992, 249–256.

[19] Kononenko I.: Estimating attributes: Analysis and extensions of Relief. L. De Raedt, & F. Bergadano (Eds.): Machine Learning: ECML-94 1994, 171–182.

[20] Kuo B. C., Ho H. H., Li C. H., Hung C. C., Taur J. S.: A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification. IEEE J Sel Top Appl Earth Obs Remote Sens 7(1)/2014, 317–326.

[21] Liu H., Yu L.: Toward integrating feature selection algorithms for classification and clustering. IEEE Transactions on knowledge and data engineering 17(4)/2005, 491–502.

[22] Neshatian K., Zhang M., Andreae P.: A filter approach to multiple feature construction for symbolic learning classifiers using genetic programming. IEEE Trans. Evol. Comput. 16(5)/2012, 645–661.

[23] Oliveira R. B., Pereira A. S., Tavaresa J. M. R. S.: Skin lesion computational diagnosis of dermoscopic images: Ensemble models based on input feature manipulation. JMRS Tavares – Computer methods and programs Computer Methods and Programs in Biomedicine 149/2017, 43–53.

[24] Oliveira R. B., Pereira A. S., Tavaresa J. M. R. S.: Skin lesion computational diagnosis of dermoscopic images: Ensemble models based on input feature manipulation. JMRS Tavares – Computer methods and programs Computer Methods and Programs in Biomedicine 149/2017, 43–53.

[25] Pal M., Foody G. M.,: Feature selection for classification of hyperspectral data by SVM. IEEE Trans Geosci Remote Sens 48(5)/2010, 2297–2307.

[26] Qi C., Zhou Z., Sun Y., Song H., Hu L., Wang Q.: Feature selection and multiple kernel boosting framework based on PSO with mutation mechanism for hyperspectral classification. Neurocomputing 220/2017, 181–190.

[27] Ramezani M, Karimian A, Moallem P.: Automatic Detection of Malignant Melanoma using Macroscopic Images. J Med Signals Sens. 4(4)/2014, 281–290.

[28] Robnik-Šikonja M., Kononenko I.: Theoretical and empirical analysis of ReliefF and RReliefF. Machine learning 53(1–2)/2003, 23–69.

[29] Sadri A. R., Azarianpour S., Zekri M., Celebi M. E., Sadri S.: WN-based approach to melanoma diagnosis from dermoscopy images. IET Image Process. 11(7)/2017, 475–482.

[30] Shahid M., Khan S.: Dermoscopy Images classification based on color, texture and shape features using SVM. The 3rd International Conference on Next Generation Computing (INC GC2017b) 2017, 243–245.

[31] Stapor K., Automatyczna klasyfikacja obiektów. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2005.

[32] UCI Machine Learning Repository [http://archive.ics.uci.edu/ml/datasets.html].

[33] Ul Ain B., Xue B., Al-Sahaf H., Zhang M.: Genetic programming for feature selection and feature construction in skin cancer image classification. Pacific Rim International Conference on Artificial Intelligence, Springer 2018, 732–745.

[34] Witten I. H., Frank E., Hall M. A.: Data mining: Practical machine learning tools and techniques. Morgan Kaufmann 2011.

[35] Xie F., Fan H., Li Y., Jiang Z., Meng R., Bovik A.: Melanoma classification on dermoscopy images using a neural network ensemble model, IEEE Transactions on Medical Imaging 36(3)/2017, 849–858.

[36] Xue B., Zhang M., Browne W. N., Yao X.: A survey on evolutionary computation approaches to feature selection. IEEE Trans. Evol. Comput. 20(4)/2016, 606–626.

[37] Yu J., Almal A. A., Dhanasekaran S. M.,Ghosh D., Worzel W. P., Chinnaiyan A., M.: Feature selection and molecular classification of cancer using genetic programming. Neoplasia 9(4)/2007, 292–303.

[38] Zagrouba E., Barhoumi W.: An accelerated system for melanoma diagnosis based on subset feature selection. Journal of Computing and Information Technology – CIT 13(1)/2005, 69–82.

[39] Zhou X., Wang J. J.: Feature selection for image classification based on a new ranking criterion. Journal of Computer and Communications 3/2015, 74–79 [http://doi.org/10.4236/jcc.2015.33013].

M.Sc. Magdalena Michalska
e-mail: magdalena.michalska@pollub.edu.pl

Ph.D. student at the Department of Electronics and Information Technology, Lublin University of Technology. Recent graduate of Warsaw University of Technology. Her research field covers medical image processing, 3D modelling, optoelectronics, and spectrophotometry. Author of more than 10 publications.

http://orcid.org/0000-0002-0874-3285

otrzymano/received: 19.01.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2577

DEVELOPING SOLUTION FOR USING ARTIFICIAL INTELLIGENCE TO OBTAIN MORE ACCURATE RESULTS OF THE BASIC PARAMETERS OF RADIO SIGNAL PROPAGATION

Andrii Shchepak, Volodimir Parkhomenko, Vyacheslav Parkhomenko
State University of Telecommunications, Kiev, Ukraine

Abstract. *The article considers the methods of calculating radio signal power. The main factors influencing the distribution and their connection with the error in the calculations of the indicators' peak values are analyzed. The regularities of signal propagation and the correlation between the distance from the radio signal source and the ratio of noise to useful information are determined. These patterns allow us to develop a model of artificial intelligence, which improves the prediction of results compared to existing calculation methods. The obtained results present the efficiency of the offered method.*

Keywords: artificial intelligence, cellular neural networks

OPRACOWANIE METODY WYKORZYSTANIA SZTUCZNEJ INTELIGENCJI DO UZYSKIWANIA DOKŁADNIEJSZYCH WYNIKÓW PODSTAWOWYCH PARAMETRÓW PROPAGACJI SYGNAŁÓW RADIOWYCH

Streszczenie. *Rozważono metody obliczania mocy sygnału radiowego. Przeanalizowano wpływ głównych czynników na rozkład i ich związek z błędem w obliczeniach wartości szczytowych wskaźników. Wyznaczono zależności na propagację sygnału i korelację między odległością od źródła sygnału radiowego a stosunkiem szumu do informacji użytecznej. Zależności te pozwalają na opracowanie modelu sztucznej inteligencji, który poprawia predykcję wyników w porównaniu z dotychczasowymi metodami obliczeń. Uzyskane wyniki pozwalają wnioskować o skuteczności oferowanej metody*.

Słowa kluczowe: sztuczna inteligencja, komórkowe sieci neuronowe

Introduction

One of the important tasks in modern radio electronics is to describe the transmission of radio waves to calculate the radiation coverage area and the quality of information exchange. The paper proposes to explore the usage of AI to generate more accurate results based on a limited amount of input data.

One of the most common ways to calculate the coverage of radio waves is to measure the input data in all necessary points. Despite all the advantages of this method, it has several significant disadvantages that in certain conditions make it impossible to use, primarily the price and the impossibility of measuring due to the unavailability of a certain area. As a replacement for this and other methods, it is proposed to consider the use of artificial intelligence and a limited amount of existing data.

1. Standard options for calculating the transmission of the radio signal

To study the features of the method, a conditional location of objects is proposed. For example, at point E is the only source of propagation of radio signals. Between A-E there is a source of deterioration of the radio signal transmission, but there is no additional data on its effect on the radio signal. Despite the existence of methods for estimating signal strength after the passage of static structures, in a limited number of measurements it is impossible to obtain the results of calculations of sufficient accuracy [4]. The gaps EB and BD are of equal length. Measurements such as altitude and radio signal strength are performed at the tops of these intervals.

Figure 1 shows the schematic location of the transmitter and measuring points of signal power and receiver height above sea level. Based on neither this information, it is proposed to calculate the data on the associated area.

The ABCD figure creates a structure that can be considered as a separate cell. Given the presence of measurements only on the vertices of the figures, an important task in calculating the propagation of the radio signal is to predict its parameters at each point. In addition to the fact that it is not possible to take into account all the extremum points on the plane in the absence of additional data, it is possible to obtain satisfactory accuracy of calculations.

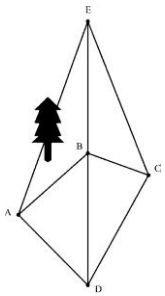


Fig. 1. Location of conditional objects

At the initial calculation of the transmission of radio waves uses software Atoll 3.1.2, which allows you to get results for all basic models with the ability to change them and the basic parameters for them. However, this, as well as other software, has a significant disadvantage – the lack of quality topographic maps for the territory of Ukraine, which ultimately has a negative impact on the quality of calculation results. The calculation data for the test environment are indicated on the graph under number.

Known methods of calculation allow to obtain the result on such parameters as the power of the radio signal with a sufficient amount of input data, but without taking into account the heterogeneity of the conditions of signal transmission in certain parts of the path.

2. Prerequisites for using AI for calculations

There are several ways to solve the problem of insufficient quantity and/or quality of input data.

Forecasting should be based on two methods:

- calculation of input data based on already known measurement results from neighboring points
- calculation of information in adjacent paths with available information about the signal quality at a certain point

Each of these methods has advantages and disadvantages. The first method allows to obtain approximate results in the spatial area in the presence of at least four points, by dividing the group of all measurement sites into cells with the union of every four adjacent points, provided that two adjacent sides of the cell do not intersect. Modern computer technology is able to quickly process

large amounts of information, so such calculations can be performed almost without delay, which is an important condition for moving points of reception/transmission. However, there are significant disadvantages – first, the calculations show low accuracy in the case of rapid or point change of conditions. Secondly, the quality of the results in the composition of the cell strongly depends on the quality of measurements at each of its vertices.

The second method allows you to calculate the average value of the parameters on the entire line between the transmitter and the point where the data of signal quality and other measurements. Among the disadvantages of this method is that the generated values do not take into account peaks and abrupt changes, but reflect only close to the average value. This approach is better than the first method in many cases in which such calculations are used. Using data based on the calculation of the average value is incorrect, so for the Earth's surface is needed to select data based on the binding to WGS 1984. The method also demonstrates different accuracy between the transmitter and the measuring point, and between the measuring point and beyond. The advantages include, firstly, the increase in the quality of measurements due to the obtained measurement results at the intermediate point. Secondly, although the obtained data do not show high accuracy in the value of signal quality at each point of the measurement interval on the segment, but allow you to accurately set the highs and lows of both signal quality and peak values of its change in real time.

To minimize the negative aspects of both methods and achieve maximum results, it is proposed to use a combination of these two methods, which is possible due to the fact that most measurements can be used for both methods. An example of the best input conditions is when two opposite cell vertices are on the same line from the transmitter, and the distance between the transmitter and the nearest vertex is equal to the distance between it and the vertex that has no direct connection with it. Among the main measurements are those that are required to achieve a high level of calculation accuracy: the height at which the receiver is located; soil height above sea level; the presence of harmful radiation at the point of measurement; the presence of large static structures or trees.

2.1. Options for using AI to calculate radio signal propagation

To obtain the results of information processing, it is proposed to build a learning model. “Supervised learning” is used to prepare the model, as the data entered must be prepared and marked in advance. Python programming language tools and Tensorflow, Keras and Pandas libraries are used to create the model.

At the first stage of the project, only the height of the soil above sea level is used. This parameter is important for further model development, as an important step will be to create a model to identify interference based on topographic data or a satellite image of the surface. Therefore, with sufficient accuracy to determine the altitude, it can be used as a basis for further calculation of the total altitude, taking into account the possible presence of stationary objects and / or trees that impair the patency of the radio signal or the reliability of the information transmission along the entire length of the path [6].

Figure 2 shows the data obtained when calculating the power of the radio signal by the formula without information from the intermediate point, and calculations of data in the presence of information from the intermediate point, as well as using the AI model.

The legend of the graph under number 3 shows the results of the calculation of the radio signal power without taking into account the data from the intermediate point. This leads to the fact that all values are average without taking into account the oscillations and deviations that occurred due to different properties of the environment during the passage of the radio signal.

The legend of the graph under numbers 1-2 demonstrates the results of calculations in the case of results from an intermediate point. In case 1, in contrast to case 2, the processing of calculation results by the basic AI model or few models in a row is additionally used [3].

The initial number of 100 positive and negative examples with good quality with known results of signal strength measurements were used to teach the model. The result of data processing on the test information is an improvement in the calculation of the radio signal power level compared to the situation when AI models are not used. The most accurate and reliable indicators of calculation are shown by the multilevel model of AI that allows to create the most exact network forecasting at the design stage.

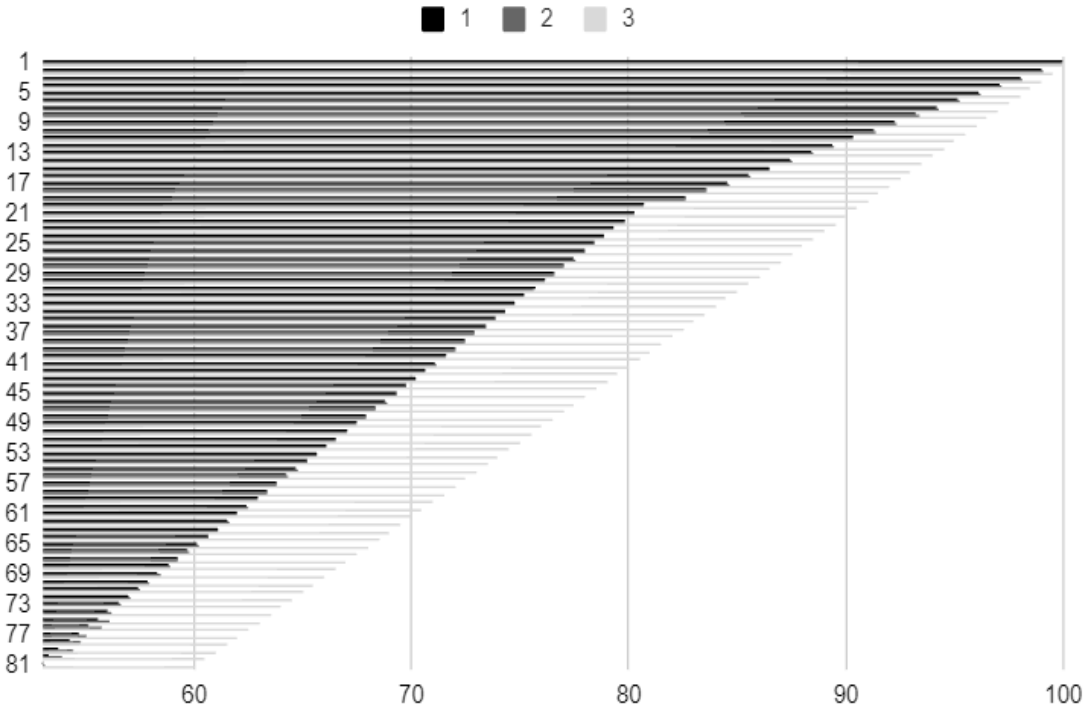


Fig. 2. Dependence of the results of calculations of radio signal power at a distance from the transmitter in relation to the maximum (100%) with three different methods of calculation with the same input data

2.2. Changes, which can be achieved by using the proposed technology

The difference between the normal calculation and the calculation using the model that was learned from the test data is shown in figure 3.

The difference in absolute terms is not significant, because the formula for calculating the power of the radio signal is quite accurate. However, this difference is significant and obvious in relative terms. Taking into account the results obtained when designing radio networks will in the future improve the transmission of radio signals, avoid the negative effects that occur during miscalculations, as well as increase the coverage area of the network.

It is also advisable to investigate the reasons why the difference between the calculated and predicted level differs the more, the greater the distance traveled by the signal from the transmitter to the place of analysis. First, it is obvious that with decreasing signal power, the increase in the ratio of the noise level to the level of useful information occurs faster, than in the start of the line.

It is important to understand that the more training cycles the model goes through and the more input data is taken into account, the more accurate the indicators will be. This is achieved due to the fact that the model replaces the mechanism of approximation of the function of calculating indicators with an extremely large number of auxiliary functions, which would not be possible without the use of such a method.

Another, more important reason is the existence of the Fresnel zones. If there is an object within a particular Fresnel zone, part of the initial signal that does not move in the direction of the antenna, and will propagate in some other direction in space, may be partially reflected from the object and fall to the receiving antenna. This reflected signal can lead to constructive or destructive interference on the receiver, depending on the path length and phase shift angle, which will result in the receiver reflected wave. The greater the distance between the transmitter and receiver, the larger the diameter of this zone, and hence the effect on the signal transmission of any object that may fall into this zone. For example, for a distance of 2.8 km, the radius of the zone can be 7.8 meters [1]. At a time when it is difficult to take such risks into account when calculating according to the formula, the use of AI models allows to significantly change the final result of calculations.

If the calculation of the diameter of the Fresnel zones is entered as an additional input indicator, it will definitely have a positive effect on the quality and accuracy of the data. However, a large amount of input data increases the learning time of the model and processing of the finished model of input indicators, and can also lead to a phenomenon known as the formation of harmful dependencies, which will ultimately reduce the quality and usefulness of the result [2].

The obtained data can be used in the calculation of the area of the radio signal. Improving the accuracy of calculating the signal power levels at the vertices of the cell will have a positive effect on the calculation of information at each of its points. Obtaining more complete and accurate data at the design stage of the network will calculate the required equipment without unnecessary information redundancy, while maximizing the reliability of message delivery and speed of information transmission, wich will help to avoid communication losses and other negative phenomena.

Important data in the design for the calculation of network coverage are the height of the transmitter, receiver above sea level, the power level of the useful signal, the presence of other signals in the area of the receiver. Taking these parameters into account when building the model, it will be possible to increase the accuracy and reliability of radio or mobile network forecasting at the design stage.

2.3. Future directions of research on the basis of the received data

A promising technology that needs to be implemented in the next stage is the recognition of satellite imagery, followed by the study of important indicators that affect the level of signal transmission by the radio path. One example of such recognition is the determination of the presence of greenery to take into account their impact on the power of the radio signal from transmitter to receiver.

In order to take into account the satellite image data in the analysis of the results, it is necessary to use separate models – one for recognizing geodata from the earth image, and the other for use in predicting the calculated information in the form of interference throughout the signal propagation. The most important task, in this case, is to determine the presence of buildings, tall trees and other static obstacles in the entire plane of the cell. Another important part is the calculation of altitude at each point, because even in the absence of static interference, the signal can be degraded by sudden changes in soil height. It also allows you to calculate the minimum height at which the receiver will be located.

This technology is promising and will allow you to build a map of coverage and radio signal distribution based on a limited amount of data, while showing high accuracy of prediction. The presence of satellite images can be replaced by images of the Earth's surface from the camera in live mode, which will allow you to calculate the generation of radio coverage of surface without delay.

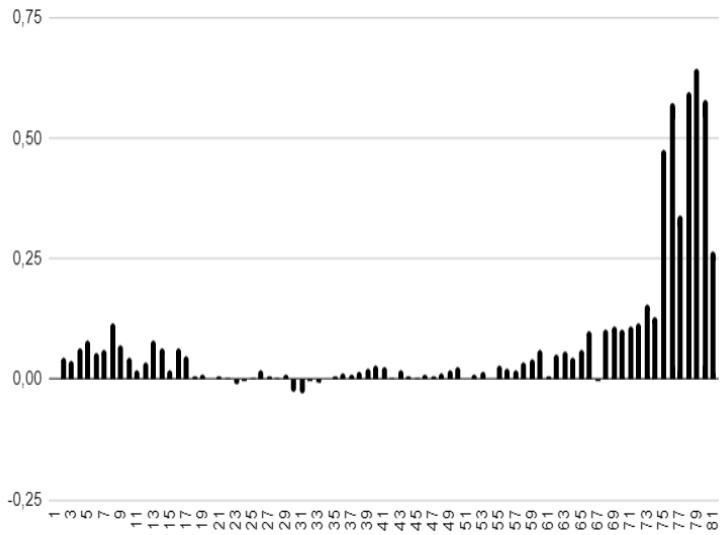


Fig. 3. The absolute difference between the peak power of the radio signal for the case without and using the AI model when calculating them relative to the distance to the transmitter

3. Conclusions

The article considers a promising technology of using AI for post-processing of input data on radio signal propagation. With the help of the trained model the level of power of a radio signal which can be received by the receiver at a certain distance from the transmitter is predicted. The use of this method together with standard calculations allows to increase the accuracy of the original data, thereby improving the final result at the stage of design and active use of the radio network or a single radio path. The results obtained in relative terms improve quality of forecasting required indicators.

Future tools for further development of the use of AI technology in the design of radio networks were also proposed. If this tool is successfully developed, it will be possible to process real-time data obtained in limited quantities, as well as images of the Earth obtained with a camera from a moving object (e.g. aircraft), and on the basis of this information to analyze the radio network.

The quality of the obtained result directly depends on the number of processed examples, but the model that has already been trained can be reused any number of times and on a variable medium, which makes it possible to analyze the environment in live mode. However, the resource requirements of the system required for data processing are not significant, and therefore the possibility of widespread use of this technology is available and may be recommended for industries such as radio network design, geodata calculation, transport network and transmission line design to avoid their negative impact.

References

[1] en.wikipedia.org/wiki/Fresnel_zone (available: 27.02.2021).
[2] en.wikipedia.org/wiki/Overfitting (available: 27.02.2021).
[3] Gafarov F. M., Halymanov A. F.: Artificial neural networks and their appendices. Kazan University Press, Kazan 2018.
[4] Miura Y., Oda Y., Taka T.: Outdoor-to indoor propagation modelling with the identification of path passing through wall openings. IEEE 13th PIMRC 1/2002, 130–134.
[5] Parkhomenko V. L.: Basics of television and radio broadcasting. State University of Telecommunications, Kyiv 2017.
[6] Rudoy G. I.: The Choice of the Activation Function in the Prediction of Neural Networks. Journal of Machine Learning and Data Analysis 1(1)/2011, 16–39 [http://jmlda.org/papers/doc/2011/no1/Rudoy2011Selection.pdf].

M.Sc. Andrii Shchepak
e-mail: k17pine@gmail.com

Graduate student of State University of Telecommunications, Kyiv, Ukraine. Research interests: telecommunications. Publications: about 3 scientific publications.

http://orcid.org/0000-0002-6570-1129

Ph.D. Volodimir Parkhomenko
e-mail: Volodymyr.P46@gmail.com

Candidate of Technical Sciences, Senior Researcher of State University of Telecommunications, Kyiv, Ukraine. Research interests: telecommunications. Publications: about 130 scientific publications.

http://orcid.org/0000-0002-7237-4330

M.Sc. Vyacheslav Parkhomenko
e-mail: viacheslav.p7788@gmail.com

Senior lecturer of the department of Telecommunications, State University of Telecommunications, Kyiv, Ukraine. Research interests: telecommunications. Publications: about 23 scientific publications.

http://orcid.org/0000-0002-9120-6842

otrzymano/received: 28.02.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2578

APPLICATION OF THE MATRIX FACTOR ANALYSIS METHOD FOR DETERMINING PARAMETERS OF THE OBJECTIVE FUNCTION FOR TRANSPORT RISK MINIMIZATION

Serhii Zabolotnii[1], Sergii Mogilei[2]
[1]Cherkasy State Business-College, Cherkasy, Ukraine, [2]Rauf Ablyazov East European University, Cherkasy, Ukraine

Abstract. *The paper regards a common transport problem with a non-classic optimization criterion to minimize transportation risks. It demonstrates that the risk parameters of the function could be found through the factor analysis method. Besides, considering that the problem contains several points of sending and delivering loads, the method is dealt with as a matrix. The research also regards the algorithm of matrix factor analysis application for determining parameters of the objective function for the problem to be solved. The survey results in a new method to construct the objective function for the optimization problem with probability parameters. It generally assists in suggesting a formal solution to such problems, foremost due to particular software.*

Keywords: factor analysis, objective function of optimization, transportation risk

ZASTOSOWANIE METODY ANALIZY WSPÓŁCZYNNIKA MACIERZOWEGO DO OKREŚLENIA PARAMETRÓW FUNKCJI CELU DLA MINIMALIZACJI RYZYKA W TRANSPORCIE

Streszczenie. *Artykuł dotyczy powszechnego problemu transportowego z nieklasycznym kryterium optymalizacji w celu zminimalizowania ryzyka transportowego. Wykazano, że parametry ryzyka takiej funkcji można wyznaczyć metodą analizy czynnikowej. Dodatkowo, ze względu na występowanie w zgłoszeniu problemu kilku punktów wysyłki i dostawy towarów, metoda ta jest prezentowana w postaci matrycy. W wyniku przeprowadzonych badań uzyskano nową metodę konstruowania funkcji celu dla postawionego problemu optymalizacji, której parametry mają charakter probabilistyczny. Pozwala to na ogólne sformalizowanie procesu rozwiązywania tego typu problemów – przede wszystkim przy pomocy specjalistycznego oprogramowania.*

Słowa kluczowe: analiza czynnikowa, funkcja optymalizacji celu, ryzyko transportowe

Introduction

Problem of minimizing transportation risks is raised quite dramatically amidst more and more intensive world economy globalization. Country borders have already become formally distinct and the whole world is getting more available to its population. In this aspect, long distance transportation becomes a norm. Although, the longer distance goods are delivered, the bigger is risk of their getting spoiled, damaged or completely ruined.

Taking all propositions into consideration, it seems quite correct to join the effort of all interested experts to investigate concrete issues of risk factors appearance and their elimination during transportations. Hence, it is important to make a quantitative calculation for the risk level, in both general and separately regarded values that are constituents of the aggregated ones. The calculations of the kind can form a solid basis for taking complex logistics solutions as for the time and means of departure, as well as of means of deliverance.

Considering contemporary experience in solving such problems, it should be mentioned that their solution is to be based on the so called transportation problem. Hence, it must be regarded in detail.

Classic (standard) transportation problem lies in determining an optimal plan for transport logistics when some load x_{ij} is to be delivered from m departure points to n points of delivery. The actual criterion of optimization is objective function S for minimizing the transportation costs [3]. Thus:

$$S = \sum_{i,j=1}^{m,n} c_{ij} x_{ij} \to \min \tag{1}$$

where by c_{ij} we mark transportation cost between i and j, stating that $i = \overline{1,m};\ j = \overline{1,n}$.

Problem like (1) appears to be explored enough [7], but, in case it is differently revealed, its solution could be quite specific. For example, instead of objective function for minimizing costs there might be offered criterion R, to minimize transportation risks [9]. Thus, the problem is as follows:

$$R = \sum_{i,j=1}^{m,n} r_{ij} x_{ij} \to \min \tag{2}$$

where by r_{ij} the level of transportation risk is shown for points i and j.

Obviously, problems (1) and (2) seem to be similar, though there is a great difference between coefficients c_{ij} and r_{ij}. Indeed, parameters c_{ij} of cost function are calculated in money units and are concrete enough, whereas parameters r_{ij} of risk function appear to be probability rendered risk events. In addition, the latter are as well aggregated risk markers and require further analysis.

Generally, there emerges a problem of constructing objective functions of optimization with probability parameters. Thus, in the aspect of optimization theory, to find a solution for the optimization problem a mere general target function is needed, or a certain analytical expression with variables and parameters. However, this theory never gives solution to calculating such parameters, and so, there are "auxiliary" scientific studies that matter, which could apply a more distinct description to methods and algorithms in finding the parameters for optimization function. It is important to mention that such methods and algorithms are distinctly correlated with essential peculiarities of objective functions, though, there exists, to some extent, a possibility to describe and apply quite universal approaches to solving this problem.

Thus, if problem (1) contains "cost as a parameter", then problem (2) suggests "probability as a parameter", and the research is fully concentrated on probability parameters. That means the research is aimed at describing universal approach to calculating probability parameters of the related objective function. The description is fully based on objective function of minimization as in (2).

Foremost, to calculate the aggregated risk indices (parameters) it is necessary to decompose them according to constituents. That means, risk factors must be found which affect the value of aggregated index. That enables to apply factor analysis method to solving problem (2) [1].

There are quite many instances of risk factors marked by v, whereas each factor is marked by k index, where $k = \overline{1,v}$. Here are three following factors ($v = 3$): emergency ($k = 1$), transport vehicle breakdown ($k = 2$) and force majeur circumstances ($k = 3$). Besides, effect (importance) of each factor on the value of aggregated index r_{ij} might differ. Considering that, marker w^k is introduced which is weight coefficient of k-factor effect. In addition, factor coefficients are normalized:

$$\sum_{k=1}^{v} w^k = 1; w^k > 0 \tag{3}$$

Condition (3) would be put differently considering that values of w^k coefficients differ as they depend on from which departure point *i* the load is transported to delivery point *j*. Thus, weight coefficients of influence are marked by $(w^k)_{ij}$. Further it will be demonstrated that in stating parameters of problem (2) they must be given in matrix, so the method proper is named as matrix factor analysis.

As for calculating concrete values for coefficient (3) a number of methods can be applied here, particularly the experts' method (method of expert estimates, or Delphi's method). It lies in analyzing approaches of prominent experts working in relevant fields [2].

On the other hand, another perspective method for investigating the problem can be group factor analysis [8, 10]. Its application is possible due to building up the hierarchy of influence factors; however, within the problem solution it is not to be applied as the solution requires only single hierarchy level. Besides, to achieve a required level of presentation, the problem should contain a larger number of risk factors.

1. The research problem statement and solution method

Thus, the research problem lies in finding parameters of objective function (2). To solve this problem, the approach suggested in [6] must be applied. The following proportion means:

$$r_{ij} = \sum_{k=1}^{v} (a^k)_{ij} \times (w^k)_{ij} + o_{ij} \tag{4}$$

where index of risk level of *k* factor is marked by a^k, and o_{ij} is the "gaussian noise", the value of which within the research problem may be neglected.

Generally, the matter of neglecting "gaussian noise" remains under consideration. Actually, the more complete is the number *v* of risk factors of transportation, the smaller this value would be.

It must be admitted, the $(a^k)_{ij}$ coefficients are as well probable for certain risk events to occur. However, in contrast with r_{ij}, they are not aggregated indices therefore it is possible to calculate their value based on results of real observations assisted by familiar methods of mathematical statistics [4].

Furthermore, sense of proportion given in expression (4) must be specified. Suppose, there are $m = 2$ points of departure and $n = 3$ delivery points. Then, for instance, for $k = 1$ the equation is:

$$(a^1)_{ij} = \begin{pmatrix} a^1_{11} & a^1_{12} & a^1_{13} \\ a^1_{21} & a^1_{22} & a^1_{23} \end{pmatrix} \tag{5}$$

In case, when weights $(w^k)_{ij}$ are given in simplified form w^k, first item in proportion (4) is a common multiplication of coefficients of matrix (5) by w^1. In the opposite case, there may be a certain relation between two matrices. With $k = 1$ the coefficient matrix $(w^k)_{ij}$ appears to be:

$$(w^1)_{ij} = \begin{pmatrix} w^1_{11} & w^1_{12} & w^1_{13} \\ w^1_{21} & w^1_{22} & w^1_{23} \end{pmatrix} \tag{6}$$

That means the relation in (4) for expressions (5) and (6) is revealed as follows:

$$(a^1)_{ij} \times (w^1)_{ij} = \begin{pmatrix} a^1_{11} \cdot w^1_{11} & a^1_{12} \cdot w^1_{12} & a^1_{13} \cdot w^1_{13} \\ a^1_{21} \cdot w^1_{21} & a^1_{22} \cdot w^1_{22} & a^1_{23} \cdot w^1_{23} \end{pmatrix} \tag{7}$$

With $k = 2$ and $k = 3$ the expressions (5) – (7) are put similarly.

For ultimate value of aggregated risk indices r_{ij}, the formula for their calculation is put for r_{11}:

$$r_{11} = a^1_{11} \cdot w^1_{11} + a^2_{11} \cdot w^2_{11} + a^3_{11} \cdot w^3_{11} \tag{8}$$

It means, to calculate parameters of objective function (2), relation (7) was additionally introduced. Therefore, the question arises if it is possible to do without introducing any of such specific relations and confine to common mathematical approaches.

It is easily noticed that in (8) r_{11} is the result of scalar multiplication of following vectors:

$$r_{11} = (a^1_{11}; a^2_{11}; a^3_{11}) \cdot (w^1_{11}; w^2_{11}; w^3_{11}) = \begin{pmatrix} a^1_{11} & a^2_{11} & a^3_{11} \end{pmatrix} \cdot \begin{pmatrix} w^1_{11} \\ w^2_{11} \\ w^3_{11} \end{pmatrix} \tag{9}$$

Expression (9) is put simultaneously for r_{11} and r_{21}:

$$\begin{pmatrix} r_{11} & 0 \\ 0 & r_{21} \end{pmatrix} = \begin{pmatrix} a^1_{11} & a^2_{11} & a^3_{11} \\ a^1_{21} & a^2_{21} & a^3_{21} \end{pmatrix} \cdot \begin{pmatrix} w^1_{11} & w^1_{21} \\ w^2_{11} & w^2_{21} \\ w^3_{11} & w^3_{21} \end{pmatrix} \tag{10}$$

Expression (10) occurs only if in case of unequal upper and lower indices in parameters $(w^k)_{ij}$ and $(a^k)_{ij}$ their multiplication is 0. This condition is to be named the equality of indices. It must be assumed that to check this condition without proper software is quite cumbersome. Therefore, it seems obvious to better solve it applying specific program instruments.

The indices being equal, expression (10) can be extended for all aggregated parameters r_{ij}:

$$\begin{pmatrix} \dots & 0 & 0 \\ 0 & r_{ij} & 0 \\ 0 & 0 & \dots \end{pmatrix} = \begin{pmatrix} & \dots & \\ a^1_{ij} & a^2_{ij} \dots & a^v_{ij} \\ & \dots & \end{pmatrix} \cdot \begin{pmatrix} & w^1_{ij} & \\ & w^2_{ij} & \\ \dots & \dots & \dots \\ & w^v_{ij} & \end{pmatrix} \tag{11}$$

Or, if to make matrix transposition $(w^k)_{ij}$ like (6):

$$\begin{pmatrix} \dots & 0 & 0 \\ 0 & r_{ij} & 0 \\ 0 & 0 & \dots \end{pmatrix} = \begin{pmatrix} & \dots & \\ a^1_{ij} & a^2_{ij} \dots & a^v_{ij} \\ & \dots & \end{pmatrix} \cdot \begin{pmatrix} & \dots & \\ w^1_{ij} & w^2_{ij} \dots & w^v_{ij} \\ & \dots & \end{pmatrix}^T \tag{12}$$

Then, relation (4) may be put as:

$$(r_{ij}) = \sum_{k=1}^{v} (a^k)_{ij} \times (w^k)_{ij} = (a^k)_{ij} \cdot (w^k)^T_{ij} \tag{13}$$

where by (r_{ij}) diagonal matrix with elements r_{ij} of main diagonal is marked.

Thus, in case parameters of objective function are calculated as probability for certain events to occur or not, it is necessary to state the factors that affect the value of these probabilities. That would enable to calculate the given parameters with the above mentioned approach.

It is to be admitted, in case the mentioned risk factors are regarded as certain aggregated values thus implying there is another level of influence factors hierarchy that enables to apply group factor analysis method to solve the issues of the problem given. In fact, that would mean another (11) – (13) iteration to calculate $(a^k)_{ij}$ coefficients. Hence, in the aspect of program solution of the problem, its solving in this case is never complicated, as all key procedures and functions are to be preliminarily taken into consideration.

The example of the research problem solution is given further with model data.

2. The research problem solution for model instance

Supposing that problem (2) contains $m = 2$ and $n = 3$, concrete values for (10) are:

$$\begin{pmatrix} r_{11} & 0 \\ 0 & r_{21} \end{pmatrix} = \begin{pmatrix} 0.1 & 0.15 & 0.12 \\ 0.2 & 0.22 & 0.16 \end{pmatrix} \cdot \begin{pmatrix} 0.5 & 0.4 \\ 0.4 & 0.3 \\ 0.1 & 0.3 \end{pmatrix} \tag{14}$$

It should be mentioned, that condition (3) for $(w^k)_{ij}$ being definitely completed. Further values are put similarly:

$$\begin{pmatrix} r_{12} & 0 \\ 0 & r_{22} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.17 & 0.05 \\ 0.1 & 0.19 & 0.24 \end{pmatrix} \cdot \begin{pmatrix} 0.3 & 0.2 \\ 0.5 & 0.7 \\ 0.2 & 0.1 \end{pmatrix} \tag{15}$$

$$\begin{pmatrix} r_{13} & 0 \\ 0 & r_{23} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.22 & 0.11 \\ 0.13 & 0.26 & 0.18 \end{pmatrix} \cdot \begin{pmatrix} 0.7 & 0.8 \\ 0.1 & 0.1 \\ 0.2 & 0.1 \end{pmatrix} \tag{16}$$

For program completion of the problem, expressions (14) – (16) can be joined into a single data array and processed by regulation (12). The condition of the indices equality being as well considered.

Hence,

$$\begin{pmatrix} r_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & r_{21} & 0 & 0 & 0 & 0 \\ 0 & 0 & r_{12} & 0 & 0 & 0 \\ 0 & 0 & 0 & r_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & r_{13} & 0 \\ 0 & 0 & 0 & 0 & 0 & r_{23} \end{pmatrix} = \tag{17}$$

$$= \begin{pmatrix} 0.1 & 0.15 & 0.12 \\ 0.2 & 0.22 & 0.16 \\ 0.3 & 0.17 & 0.05 \\ 0.1 & 0.19 & 0.24 \\ 0.2 & 0.22 & 0.11 \\ 0.13 & 0.26 & 0.18 \end{pmatrix} \cdot \begin{pmatrix} 0.5 & 0.4 & 0.3 & 0.2 & 0.7 & 0.8 \\ 0.4 & 0.3 & 0.5 & 0.7 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.2 & 0.1 & 0.2 & 0.1 \end{pmatrix}$$

The simple calculations would result in:

$$\begin{pmatrix} r_{11} & 0 \\ 0 & r_{21} \end{pmatrix} =$$

$$= \begin{pmatrix} 0.1 \cdot 0.5 + 0.15 \cdot 0.4 + 0.12 \cdot 0.1 & 0 \\ 0 & 0.2 \cdot 0.4 + 0.22 \cdot 0.3 + 0.16 \cdot 03 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.05 + 0.06 + 0.012 & 0 \\ 0 & 0.08 + 0.066 + 0.048 \end{pmatrix}$$

$$\begin{pmatrix} r_{12} & 0 \\ 0 & r_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} 0.3 \cdot 0.3 + 0.17 \cdot 0.5 + 0.05 \cdot 0.2 & 0 \\ 0 & 0.1 \cdot 0.2 + 0.19 \cdot 0.7 + 0.24 \cdot 0.1 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.09 + 0.085 + 0.01 & 0 \\ 0 & 0.02 + 0.133 + 0.024 \end{pmatrix}$$

$$\begin{pmatrix} r_{13} & 0 \\ 0 & r_{23} \end{pmatrix} =$$

$$= \begin{pmatrix} 0.2 \cdot 0.7 + 0.22 \cdot 0.1 + 0.11 \cdot 0.2 & 0 \\ 0 & 0.13 \cdot 0.8 + 0.26 \cdot 0.1 + 0.18 \cdot 0.1 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.14 + 0.022 + 0.022 & 0 \\ 0 & 0.104 + 0.026 + 0.018 \end{pmatrix}$$

Ultimately,

$$\begin{pmatrix} r_{11} & 0 \\ 0 & r_{21} \end{pmatrix} = \begin{pmatrix} 0.122 & 0 \\ 0 & 0.194 \end{pmatrix} \tag{18}$$

$$\begin{pmatrix} r_{12} & 0 \\ 0 & r_{22} \end{pmatrix} = \begin{pmatrix} 0.185 & 0 \\ 0 & 0.177 \end{pmatrix} \tag{19}$$

$$\begin{pmatrix} r_{13} & 0 \\ 0 & r_{23} \end{pmatrix} = \begin{pmatrix} 0.184 & 0 \\ 0 & 0.148 \end{pmatrix} \tag{20}$$

or as (17):

$$\begin{pmatrix} r_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & r_{21} & 0 & 0 & 0 & 0 \\ 0 & 0 & r_{12} & 0 & 0 & 0 \\ 0 & 0 & 0 & r_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & r_{13} & 0 \\ 0 & 0 & 0 & 0 & 0 & r_{23} \end{pmatrix} = \tag{21}$$

$$= \begin{pmatrix} 0.122 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.194 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.185 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.177 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.184 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.148 \end{pmatrix}$$

Results (18) – (20) are checked by MS Excel. Table 1 contains values of $(a^k)_{ij}$ coefficients.

Table 1. Indices of risk level for each factor

$(a^1)_{ij}$	$(a^2)_{ij}$	$(a^3)_{ij}$
0.1	0.15	0.12
0.2	0.22	0.16
0.3	0.17	0.05
0.1	0.19	0.24
0.2	0.22	0.11
0.13	0.26	0.18

Table 2 contains values of $(w^k)_{ij}$ coefficients.

Table 2. Indices of weight for risk factors

$(w^1)_{ij}$	$(w^2)_{ij}$	$(w^3)_{ij}$
0.5	0.4	0.1
0.4	0.3	0.3
0.3	0.5	0.2
0.2	0.7	0.1
0.7	0.1	0.2
0.8	0.1	0.1

Results of calculations are shown in Figure 1.

0.122	0.000	0.000	0.00	0.00	0.00
0.000	0.194	0.000	0.00	0.00	0.00
0.000	0.000	0.185	0.000	0.000	0.000
0.000	0.000	0.000	0.177	0.000	0.000
0.000	0.000	0.000	0.000	0.184	0.000
0.000	0.000	0.000	0.000	0.000	0.148

Fig. 1. Results of calculations by MS Excel

It must be admitted, MS Excel is not at all the only program mathematical means by which the solution is possible, other specialized instruments alike being Mathcad and Matlab [5]. However, it is clear that the condition of indices equality can't be checked by those program means. The ultimate result of problem solution is given as a row vector:

(r_{ij}) = (0.122; 0.194; 0.185; 0.177; 0.184; 0.148).

Obviously, results of handmade calculations coincide with those obtained by applied software. A more distinct computer solution of the problem could be performed by open code programs, through the condition of indices equality check.

3. Conclusions

The paper regards classical transport problem which presupposes calculation of optimal plan of transportations from several departure points to several points of delivery. Instead of criterion of minimal transportation cost as objective function, minimization of their risk level is taken. Values of parameters for risk function of optimization appear to have a sense of probability. To calculate them, method of factor analysis is applied as a matrix that is matrix factor analysis.

Model example demonstrates the work of the method, the calculations of objective function parameters being made both by hand and by MS Excel. The results of problem solution are shown to be quite similar for both cases.

In further studies, the matrix factor analysis method may be applied for solving other problems alike. It could be also regarded as algorithm basis for solving similar problems through specific software.

Another scientific interest is to be presented by similar studies in context of multimodal and intermodal transportations. The transportation problem in this case would involve far more numerous risk factors that may permit to present a more distinct perspective of the above mentioned method for creating its objective function. In addition, solving multi criteria transportation problems would permit to find further "interaction" in several different optimization criteria both probable and non-probable. As for the program solution of the problem, the above mentioned cases are hardly able to provide for ready program products in order to solve this task. Therefore, it would be necessary to apply the open-code software.

References

[1] Boyd K. C.: Factor analysis. In The Routledge Handbook of Research Methods in the Study of Religion. Taylor and Francis, 2013, 204–216 [http://doi.org/10.4324/9780203154281-22].

[2] Chalmers J., Armour M.: The Delphi technique. In Handbook of Research Methods in Health Social Sciences. Springer, Singapore 2019, 715–735 [http://doi.org/10.1007/978-981-10-5251-4_99].

[3] Díaz-Parra O., Ruiz-Vanoye J. A., Bernábe Loranca B., Fuentes-Penna A., Barrera-Cámara R.A.: A survey of transportation problems. Journal of Applied Mathematics. Hindawi Publishing Corporation 2014, 848129 [http://doi.org/10.1155/2014/848129].

[4] Freund R. J., Wilson W. J., Mohr D. L.: Data and Statistics. In Statistical Methods, 2010, 1–65 [http://doi.org/10.1016/b978-0-12-374970-3.00001-9].

[5] Honcharov A., Mogilei S.: Solving multimodal transportation problems by different program means. Bulletin of Cherkasy State Technological University 3/2020, 67–74.

[6] Klami A., Virtanen S., Leppaaho E., Kaski S.: Group Factor Analysis. IEEE Transactions on Neural Networks and Learning Systems 26(9)/2015, 2136–2147 [http://doi.org/10.1109/TNNLS.2014.2376974].

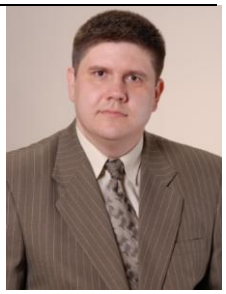
[7] Privarova R.: Operational analysis tools in solving transport task. Perner's contacts XI(2)/2016, 82–89.

[8] Virtanen S., Klami A., Khan S. A., Kaski, S.: Bayesian group factor analysis. Proceedings of the 15th AISTATS, JMLR W&CP 22/2012, 1269–1277.

[9] Zabolotnii S., Mogilei S.: The methods for determining the parameters of the objective function of multimodal transportation risk. Proceedings of V International Scientific-Practical Conference "ITEST-2020", Cherkasy 2020, 114–115.

[10] Zhao S., Gao C., Mukherjee S., Engelhardt B. E.: Bayesian group factor analysis with structured sparsity. Journal of Machine Learning Research 17/2016, 1–47.

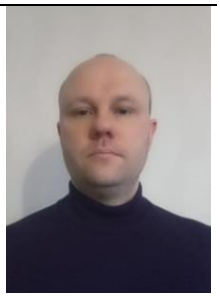
D.Sc. Serhii Zabolotnii
e-mail: zabolotniua@gmail.com



Serhii Zabolotnii was born in 1973 in Cherkasy, Ukraine. In 1995 he graduated the Department of informational technologies of Cherkasy State Technological University. Doctor of Technical Sciences (2015).
Job: professor of the Department of Computer Engineering and Informational Technologies in Cherkasy State Business-College. Scientific interests: statistical data processing, computer modeling.
Author/co-author of more than 150 publications, 1 monography and 7 patents.

http://orcid.org/0000-0003-0242-2234

M.Sc. Sergii Mogilei
e-mail: sergiymogiley@gmail.com



Sergii Mogilei is a Ph.D. student in Cherkasy State Technological University and works as a tutor in Rauf Ablyazov East European University (Cherkasy, Ukraine). His main scientific interests include mathematical and computer modeling, business analysis, mathematical programming, decision making theory and logistics.

http://orcid.org/0000-0002-9296-6827

otrzymano/received: 28.02.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2591

DESCRIPTION OF ALGORITHMS FOR BALANCING NUMERICAL MATRICES AND THEIR DIVISION INTO HIERARCHICAL LEVELS ACCORDING TO THEIR TYPE AND COMPLEXITY

Yuriy Khanas[1], Michał Borecki[2]
[1]Lviv Polytechnic National University, Lviv, Ukraine, [2]Warsaw University of Technology, Warsaw, Poland

Abstract. *This article describes a set of algorithms for so-called balancing of numerical matrices, which were developed by the author. Each section consists of several algorithms that are divided into different levels. The order of these levels depended on the chronology of the creation of certain algorithms. Chronology also affected the complexity of these balancing algorithms, so it can be argued that the algorithms are described in order from the simplest level to the most complex. It is important to emphasize that the purpose of the article is to describe the actions on matrices that determine the balancing algorithm of a certain level, and practical application will be the next step.*

Keywords: matrix balancing, separated matrix sectors, solid matrix sector, virtual matrix boundary, non-uniform matrix

ALGORYTMY BILANSOWANIA ORAZ HIERARCHIZACJI MACIERZY WEDŁUG ICH TYPU I ZŁOŻONOŚCI

Streszczenie. *Artykuł przedstawia autorskie algorytmy bilansowania macierzy. Każdy rozdział składa się z kilku algorytmów, które są rozdzielone na różne poziomy. Te poziomy są uporządkowane w zależności od chronologii ich stworzenia. Podobnie chronologia ma wpływ na złożoność algorytmów zbilansowania, w związku z tym można stwierdzać, że algorytmy są uszeregowana według stopnia złożoności. Niniejszy artykuł jest pierwszym etapem pokazującym sposób zbilansowania pewnego poziomu macierzy, natomiast kolejnym etapem będzie efekt praktyczny.*

Słowa kluczowe: bilansowanie macierzy, rozdzielone sektory macierzy, cały sektor macierzy, granica wirtualna macierzy, wyznaczona granica macierzy

Introduction

Some of the described algorithms have been mentioned earlier in previous articles of the author [1–7], but some names, designations, definitions, steps and actions may be partially or completely different, due to constant research and development of this "family" of algorithms.

The very first steps in the development of the author were aimed at creating algorithms for data compression [8], initially experiments were performed with text compression. In the further work manipulations with alphabets, keys, the size of the text and other aspects which influenced formation of a numerical matrix were carried out. After that, the obtained matrices of numbers were changed by various manipulations, the size of the matrix and the content were changed. Thus, the author came to the possibility of reducing numerical matrices [9–13]. At the beginning of matrix reduction studies, the results of each iteration had quite unpredictable consequences for the size of these matrices, for example, the number of columns could only decrease, and the number of rows could both decrease and increase. Then the author focused on methodological support, as a result of which various rules, remarks and something like theorems were developed. As experiments on matrix reduction options continued, several interesting cases arose – namely, matrices that could be completely reduced and disappeared. Matrices that were not subject to reduction were also identified and it was impossible to perform
at least one iteration. At first glance, this was a problem because such a matrix could not be reduced or restored, it was then that the author developed so-called "auxiliary" algorithms that allowed to change the contents of the matrix so that at least one iteration of reduction was allowed, but it was not reduced completely. All developments and modifications of further algorithms were accompanied by methodological support, which in turn grew more and more. Eventually, a compression algorithm was developed to reduce the amount of text data and return it to its original form.

Testing of the already developed compression method showed quite different results. In some cases, the new algorithm was more efficient than existing counterparts, and sometimes less efficient. Even the most unsuccessful experiments did not disappoint the author, but on the contrary prompted to invent a new approach, algorithm, rule, exclusion from the rules or a new key system for compressing and restoring text.

The author's plans were not limited to test information, it was the turn of research on image compression algorithms, signals, and even the development of compression algorithms with data encryption. However, in the study of data encryption algorithms, the idea arose to apply this method to encryption, and postpone compression for later.

When adapting the algorithm for data encryption, there was no need to reduce the matrices, then it was decided to take as a basis auxiliary algorithms, which the author invented earlier. The first experiments showed quite good results, even most of the methodological support remained relevant. In fact, then the algorithm moved in a new direction and developed rapidly. Through modifications, finding possible problems, the number of algorithms has grown. In fact, then there was a need to systematize them so that it was possible to clearly present the method of encryption in publications and other scientific papers.

The first sorting of all variants of algorithms had the following structure: basic algorithms, auxiliary, modified, combined, additional. However, it is difficult to lay distinguish between modified and combined algorithms, as between the subsidiary and additional, as are algorithms that can belong to multiple categories.

To avoid confusion in further work on the development and improvement of this method, the author proposed to divide these algorithms into levels, which may include several balancing algorithms.

Although this algorithm is effectively used to encrypt text, it is not a limitation for it, because at this stage several areas of science have been found where its application has the right to life.

Once again, I would like to note that in this article the emphasis is on the description of actions on matrices actually balancing algorithms. Some auxiliary algorithms have already been mentioned by the author in previous articles, such as partitioning and reflection of the matrix [11,12]. Here the focus is on balancing and dividing it into different levels, each of which will be described in its own section of the article. This article is an introduction and the basis for subsequent new articles to gradually introduce fellow scientists and other readers of this work into the field of its development, without confusing anyone.

The author emphasizes that each individual algorithm and method as a whole is constantly being developed, modified, improved, constant research is conducted, but any changes will be described in subsequent articles.

It can also be noted that the method is a kind of innovation of the author (so far in the field of data encryption), respectively, the actions, steps and manipulations shown may or may not coincide with some algorithms and mathematical operations.

1. First level balancing algorithms

The first level includes two algorithms:

- horizontal balancing,
- vertical balancing.

In the initial stages of research, they were called basic algorithms because they are the basis for the algorithms that will be described below, i.e. all modifications and combinations originate from them.

In this article, these algorithms will have a slightly more complex description and additional definitions that will allow you to more clearly see the difference between these levels. These definitions will be used in future works [12–22].

Each of the balancing types is designed for certain proportions of the matrix. Horizontal balancing is designed for a matrix that has an even number of columns and an odd number of rows. Vertical balancing, in turn, is designed for matrices with an odd number of columns and an even number of rows.

Consider an example with matrices for each of the cases, the matrices are formed from completely random numbers.

				4	0	3	
2	1	7	4	7	7	4	
4	7	7	0	1	7	8	
1	8	4	3	2	4	1	
		A				B	

Fig. 1. Starting matrices for balancing the first level

Two simple matrices for convenient and clear presentation of algorithm steps.

Accordingly, Fig. 1a – for horizontal balancing, Fig. 1b – for vertical balancing.

To demonstrate how each balancing occurs, each matrix should be divided into two equal parts, i.e. so that each part has the same number of digits. The so-called "virtual boundary" will be used for this purpose (in earlier works it was called simply the center of a matrix).

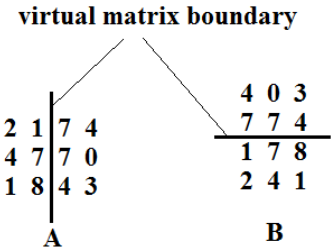


Fig. 2. Separation of matrices by a virtual boundary

When the matrices are separated by a boundary (center), then the opposite parts of the matrix are compared. In the matrix for horizontal balancing (Fig. 2a) the opposite rows will be compared, and in the matrix for vertical balancing the opposite columns will be compared. The sum of the numbers of a row or column separated by a boundary will be compared. This is called a sector-by-sector comparison.

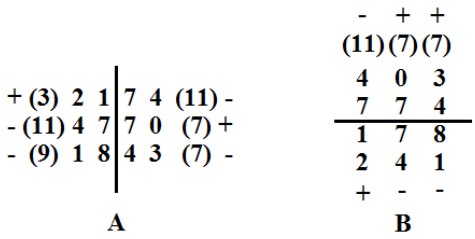


Fig. 3. Comparison of individual sectors of the matrix

The results of the comparison of sectors are shown in the figure, in parentheses are the sums of opposite rows and columns. Accordingly, the row or column with the larger sum should be reduced, and the sector with the smaller sum should be increased. This operation is as follows – when you increase the sector, all digits (row or column) increase by one, and when you decrease all the digits of the sector are reduced by one. The figure also shows the actions of increase and decrease (plus and minus).

Now we will iterate the balancing of the first level on both matrices.

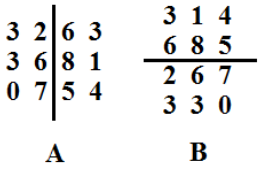


Fig. 4. Matrices after balancing the first level

Since matrices with the same numbers were chosen for both balances, the result is the same in both. At first glance, this is a primitive algorithm with simple actions and steps, but we should not forget that this is only the first algorithm and it is designed to provide an accessible explanation of the logic of the developed method.

Some publications have already described the remark about the cyclic change of matrix numbers, but it will be useful to recall it. The matrix uses unique, positive, integers (0–9), so to preserve the integrity of the matrix, the author introduced a rule: if when reducing the sector there is a number zero, then after reducing it by one, zero will become the number nine, and vice versa, if the number nine is present when the sector increases, then it becomes zero. That is, the numbers will change cyclically in the range 0–9.

However, the author has developed separate algorithms for working with fractional, multi-digit and negative numbers, but they have applications in another area, so they will not be described in this article.

There are also cases when the matrix has an even number of rows and columns, then depending on their number you can determine which type of balancing will be optimal. And if the matrix has an even number of rows and columns and their number is the same. In this case, the user can choose the type of balancing at random, or conduct each balancing separately and compare the results, and then based on them to make your choice.

Consider ode from such examples using a small matrix. Hereinafter, such matrices will be called paired symmetric (uniform) matrices.

3	5	5	4	7	0
2	7	9	0	1	3
1	4	1	1	9	8
5	4	1	9	0	0
1	8	6	3	1	8
5	5	1	3	2	0

Fig. 5. Paired symmetric matrix

Now we need to give a virtual limit, and in this example we give two – one for each type of balancing, and then compare the results.

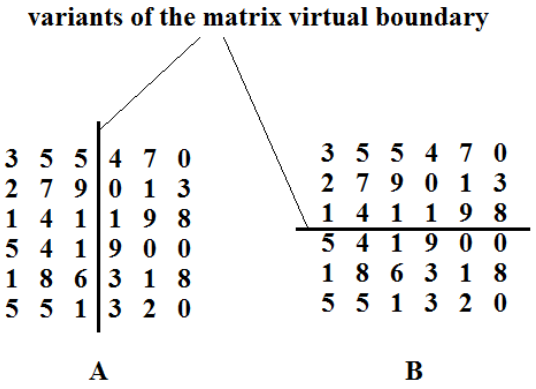


Fig. 6. Virtual boundaries of the matrix for balancing both types

The principle remains the same, comparing by divided sectors, determining the increase and decrease for each of the sectors.

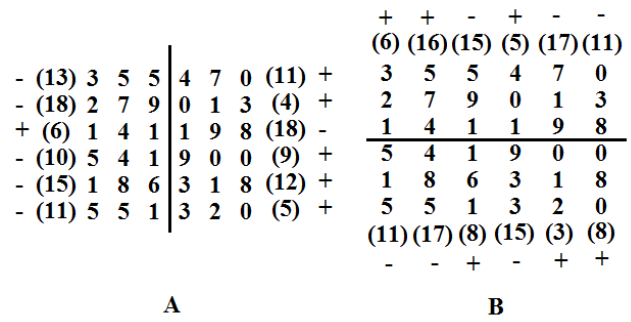


Fig. 7. Comparison of individual sectors of the matrix

Even after comparing the sectors, it can be seen that the number of sectors for increasing and decreasing differs depending on the boundary of the matrix, ie it is immediately apparent that the results of horizontal and vertical balancing will be different from each other, despite the fact that they are the same matrix.

Let's carry out one iteration of each balancing.

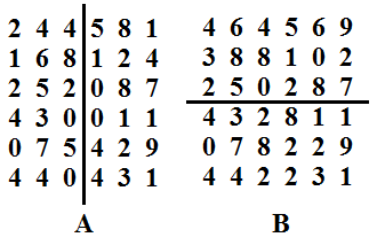


Fig. 8. The results of balancing the matrix

In general, the results of balancing the matrix differ, although some balanced numbers are the same in both cases. This fact does not indicate the ineffectiveness of any type of balancing. Each of the two types is effective, but to explore it in more detail, it would be necessary to compare not only numerical matrices, but also, for example, encrypted texts. Then you can judge which of the types of balancing will be optimal for this example. For different texts or other data, or even other scientific problems, any of the algorithms described in the article may be the most effective.

2. Second level balancing algorithms

Among the algorithms of the second level there are also our old acquaintances, namely horizontal and vertical balancing, but with a small difference – in the previous section there was a virtual boundary that divided the matrix, and at this level we have a certain limit of matrix division.

The defined boundary (center) of the matrix is the central row or column of the matrix, which will not change (its numbers will be neither increased nor decreased, but will simply remain unchanged).

The question arises – for which matrices is the balance of the second level? The answer is quite obvious: balancing of this level is intended for matrices with an odd number of rows and columns.

This section will cover a number of examples, namely examples with different ratios of an odd number of rows and columns, with both types of balancing applied to each example to illustrate the results. Also, both types of balancing will be used for a matrix with an odd number and the same number of rows and columns. Such matrices will hereinafter be called odd symmetric (uniform) matrices.

For the first example, a matrix is chosen where the number of rows is greater than the number of columns (and their number is odd).

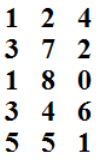


Fig. 9. Input matrix for balancing

This matrix has 5 rows and 3 columns, in both cases we can select the border in the form of a central row or a central column of the matrix. Therefore, we select both boundaries of the matrix and perform both types of balancing.

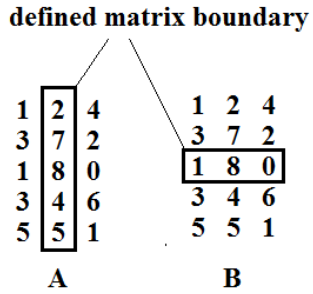


Fig. 10. Matrix with defined boundaries for horizontal and vertical balancing

Next, the standard steps for the algorithms are performed, namely the comparison of opposite sectors and the definition of sectors to increase and decrease. Previously, this process was called determining the direction of balancing, but this definition will be more logical.

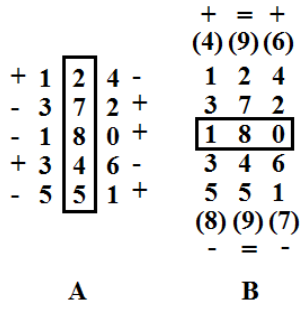


Fig. 11. Comparison of sectors to increase and decrease them

In the first version (Fig. 11a). For horizontal balancing the sum of numbers in the sectors were not written, because in each of the opposite sectors there is only one digit. In the second variant (Fig. 11b). There is a more interesting case – there are two opposite sectors of the matrix with the same sums, they are denoted by an equal sign. In some previous articles, such cases also occurred, then the user could choose which of the sectors to increase and which to decrease. In this case, we will leave these sectors unchanged. This example is worth remembering, as it will be involved in next-level algorithms, where this kind of situation will be relevant and will be resolved.

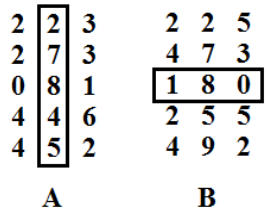


Fig. 12. Balancing of odd non-uniform matrices

For this rather simple example, the first option (horizontal balancing) looks more effective than the second (vertical balancing). This conclusion can be reached due to a larger number of balanced numbers. Because the second option had two sectors with the same sums of numbers, they were not balanced. Such cases will not happen more than once and for each case the author has developed modifications of algorithms, which will be shown in higher-level algorithms.

Now let's try this level algorithm for a matrix with more columns and fewer rows (in both cases their number will be odd).

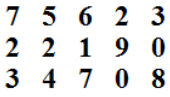


Fig. 13. Input matrix for balancing

An odd asymmetric matrix is obtained again, except that this time we have 5 columns and three rows.

We will draw certain boundaries for horizontal and vertical balancing.

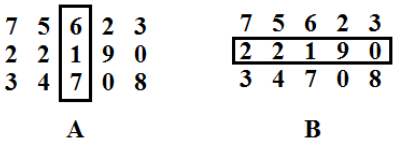


Fig. 14. Defining the boundaries of the matrix

Now we redefine the sums of the numbers in the opposite sectors of the matrix, then increase and decrease the corresponding sectors. Let's not forget that a certain limit is not subject to balancing, so it will remain unchanged.

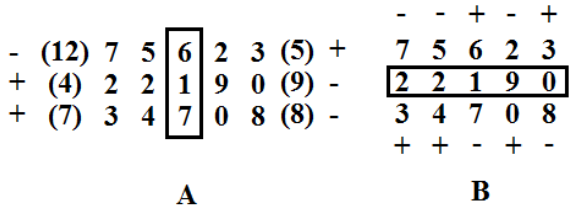


Fig. 15. The results of comparing opposite sectors

This time there are no opposite sectors with the same sums of numbers, so all sectors will be balanced. In the variant for vertical balancing (Fig. 14b.), The amounts did not fit, as each opposite sector consists of only one digit.

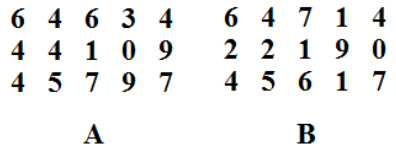


Fig. 16. Horizontally and vertically balanced matrix

As a result, both versions of the matrix were balanced without any difficulty. Of course, some figures in the balanced sectors turned out to be the same in both cases, but in general the result is different. This is due to the small size of the matrix, but in such a simple example it is easier to explain all the steps in an accessible way.

The last example in this section is the balancing of an odd uniform (symmetric) matrix in both ways.

For example, take a matrix of size 5x5.

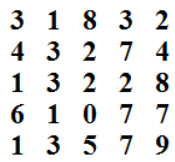


Fig. 17. Uniform odd matrix

This matrix will also be balanced horizontally and vertically, after which we compare the results.

So let's start with the defined boundaries of the matrix.

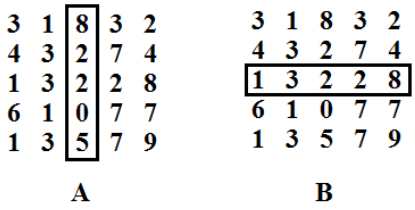


Fig. 18. The limits of the matrix for the two balancing options are defined

Next, we again compare the sectors of the matrix with each other, write the sums of the numbers of sectors and indicate which of the sectors will be increased and which will be reduced.

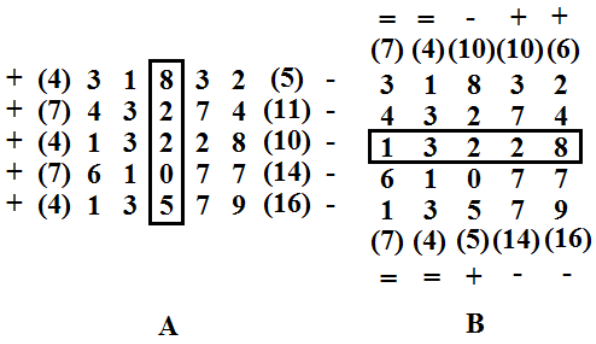


Fig. 19. Comparison of sectors of an odd symmetric matrix

In the matrix for horizontal balancing (Fig. 18a.) It is seen that all sectors on the left will be increased, and all sectors on the right will be reduced. Such cases may occur from time to time, but they do not constitute any obstacle to obtaining correct results.

Instead, the matrix for vertical balancing has four sectors (two pairs of opposite sectors) that are equal, ie the sum of the numbers of these sectors is the same. We will not balance such sectors (as in the previous similar example), ie they will remain unchanged, as well as the defined limit of the matrix.

4	2	8	2	1		3	1	7	4	3
5	4	2	6	3		4	3	1	8	5
2	4	2	1	7		1	3	2	2	8
7	1	0	6	6		6	1	1	6	6
2	4	5	6	8		1	3	6	6	8
		A						B		

Fig. 20. The results of balancing an odd uniform matrix

It is immediately clear that horizontal balancing for this variant of the matrix is more efficient than vertical, due to the sectors of the matrix, which were equal in vertical balancing. Therefore, in the first case, all the numbers were balanced except for a certain limit. Although when using combined types of balancing, one of the algorithms used may give an ineffective result, but if we take into account the total results of all involved algorithms, then the applied method will give high results (this is tested in encryption problems).

3. Third level balancing algorithms

The third level balancing algorithms contain some steps from the previous levels and several features of their modification. Here, too, the balancing algorithm is performed horizontally and vertically. In addition, both the virtual boundary of the matrix and the defined one will be taken into account. The peculiarity of the third level balancing is that the separated sectors of the matrix will no longer be compared, but continuous opposite sectors. This means that the adjacent rows or columns will not be compared, but all the digits of both halves of the matrix separated by a border (center). In this way, there will be no more cases when several separate sectors will be equal to each other. The algorithm can be used for matrices of any proportions: even, odd, symmetric, non-uniform matrices. That is, the number of rows or columns will not limit the capabilities of the algorithm.

The most obvious example is a paired symmetric matrix, which will be balanced horizontally and vertically, after which it will be possible to compare the results.

1	8	4	6	4	0		1	8	4	6	4	0
3	1	2	4	0	1		3	1	2	4	0	1
6	7	8	3	3	3		6	7	8	3	3	3
2	3	1	6	2	2		2	3	1	6	2	2
7	4	5	9	1	0		7	4	5	9	1	0
4	1	9	0	9	1		4	1	9	0	9	1
		A							B			

Fig. 21. Location of the matrix

Unlike the previous level algorithms, all numbers of the matrix sector on each side of the boundary will be added. It is assumed that absolutely all numbers on each side of the matrix will be counted as one continuous sector. This can be seen in more detail in the example.

													(64) +				
- (76)	1	8	4	6	4	0	(54) +	1	8	4	6	4	0				
	3	1	2	4	0	1		3	1	2	4	0	1				
	6	7	8	3	3	3		6	7	8	3	3	3				
	2	3	1	6	2	2		2	3	1	6	2	2				
	7	4	5	9	1	0		7	4	5	9	1	0				
	4	1	9	0	9	1		4	1	9	0	9	1				
			A							B				(66) -			

Fig. 22. Comparison of matrix sectors before balancing

For an illustrative example, the sectors of the matrix, which are separated by a boundary, were identified. Numbers in parentheses are the sum of all numbers in a sector. Accordingly, the plus and minus signs mean the increase and decrease of sectors, namely all numbers of a continuous sector.

0	7	3	7	5	1		2	9	5	7	3	1
2	0	1	5	1	2		4	2	3	5	1	2
5	6	7	4	4	4		7	8	9	4	4	4
1	2	0	7	3	3		1	2	0	5	1	1
6	3	4	0	2	1		6	3	4	8	0	9
3	0	8	1	0	2		3	0	8	9	8	0
		A							B			

Fig. 23. The matrix is balanced horizontally and vertically

The figure shows that the variants of the matrix after horizontal and vertical balancing differ from each other by only half. This is not a flaw or error of the algorithm. Such cases occur in algorithms of this level when they are used for paired uniform (symmetric matrices).

Even if you replace the virtual boundary of the matrix with a defined one, the result will be the same, the matrices after both types of balancing will differ by exactly half, except for the invariant boundary in the form of a row or column of the matrix. Therefore, there is actually no point in conducting an additional experiment, especially if the main result of the demonstration is known in advance.

Several more combined algorithms are based on this algorithm, but they belong to algorithms of another (higher level), so they will not be described in this article.

These three levels of balancing algorithms that were shown in the article are just the tip of the iceberg, in fact there are many more. In this paper, the algorithms of the initial level are described, and in the following publications the algorithms of other (higher) levels will be shown and described.

It should also be noted that each algorithm of each level is unique, useful and efficient, despite its simplicity.

4. Conclusions

This paper presents three levels of algorithms for balancing numerical matrices. These levels belong to the class of low (initial or basic) levels, because all higher levels in one way or another originate (basis) from them or derivative algorithms.

The uniqueness of this development is that with the right combination of these algorithms, we get a new and effective method of encryption (other areas of science are not excluded). Some algorithms and their combinations have proven themselves at a high level in experiments with data encryption. This cipher is almost impossible to break, because it does not match the generally accepted rules or algorithms. It is known that ideal methods, algorithms or systems do not exist in nature. However, this method is not amenable to statistical or frequency analysis and any patterns in it can not be detected.

After describing all the classes, each of which consists of several levels, which in turn consist of sets of algorithms and their variations, the author's further plans include the development of this field of science. Publications are planned, which will show and describe the application of these algorithms in practice (so far for encryption), search for combinations of developed algorithms, description of methodological support (rules, comments, definitions of abbreviations), demonstration of inverse (inverse) algorithms for decoding, presentation of mathematical formulas and mathematical models of the method, the use of special markers for the formation of encryption keys.

References

[1] Agarwal S.: Symmetric Key Encryption using Iterated Fractal Functions. International Journal of Computer Network and Information Security 9(4)/2019, 1–9.

[2] Anisimov A. V., Kulyabko P. P.: Information systems and databases: A textbook for students of the faculty computer science and cybernetics. Kyiv 2017.

[3] Bogdanov A., Khovratovich D., Rechberger D.: Biclique Cryptanalysis of the Full AES. Advances in Cryptology – ASIACRYPT 2011. Lecture Notes in Computer Science 7073. Springer, Berlin 2011.

[4] Borecki M.: Risk level analysis in the selected (initial) stage of the project life cycle. Management and production engineering review 11(4)/2020, 104–112.

[5] Borecki M., Ciuba M., Kharchenko Y., Khanas Y.: Main aspects influencing the evaluation of atmospheric overvoltages in high-voltage networks. Bull. Pol. Ac.: Tech 69(1)/ 2021, 1–8.

[6] Buryachok V. L.: The choice of a rational method of generating passwords among many existing. Information security 25(1)/2019, 59–64.

[7] Buryachok V. L.: Generate a password for wireless networks using variable rule of complication. Information protection 21(1)/2019, 52–59.

[8] Hughes J., Cybenko G.: Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity. Technology Innovation Management Review 2013, 15–24.

[9] Isa M. A. M., Hashim H., Ab Manan J. L., Adnan S. F. S., Mhmod R.: RF simulator for cryptographic protocol. IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2014, 518–523.

[10] Josefsson S., Leonard S.: Textual Encodings of PKIX, PKCS, and CMS Structures. Internet Engineering Task Force April 2015.

[11] Khanas, Y., Borecki, M.: Research on the use of algorithms for matrix transformations for encrypting text information. Security and Privacy 3(6)/2020, 1–13.

[12] Khanas Y., Ivanciv R., Litvinko S.: The algorithm for minimizing matrices in the given direction of reduction and the rules for their restoration. Visnyk of the National University "Lviv Polytechnic" 882/2017, 12–17.

[13] Khanas Y., Ivanciv R.: Application Mirroring of Matrices to Prevent Excessive Reduction. Perspective technologies and design methods of MEMST (MEMSTECH 2016), Lviv-Polyana 2016, 143–145.

[14] Krasilenko V. G.: Multifunctional parametric matrix-algebraic models (MAM) of cryptographic transformations (CP) with modular operations and their modeling. 72 NPK – conference materials, Odessa 2017, 123–128.

[15] Krasilenko V. G.: Improvement and modeling of electronic digital signatures of matrix type for textographic documents. Proceedings of the VI International Scientific and Practical Conference "Information Control Systems and Technologies" (IUST-Odessa-2017), Odessa 2017.

[16] Krasylenko V. G., Nikitovich D. V., Yatskovskaya R. O., Yatskovsky V. I.: Simulation of advanced multi-step 2D RSA algorithms for cryptographic transformations and blind electronic digital signature. Information processing systems 1(156)/2019, 92–100.

[17] Leurent G., Peyrin T.: SHA-1 is a Shambles. First Chosen-Prefix Collision on SHA-1 and Application to the PGP Web of Trust. Real World Crypto 2020.

[18] Lobur M. V., Ivantsiv R. D., Kolesnyk K. K., Khanas Y. Y.: Development of an algorithm for reducing matrices depending on their size and content. Scientific and Technical Journal "Instrumentation Technology" 2/2016, 29–31.

[19] Nazarkevich M. A., Dronyuk I. M., Troyan O. A., Tomashchuk T. Yu.: Development of a method of document protection by latent elements based on fractals. Information protection 17(1)/2015, 21–26.

[20] Nikonov V. G., Zobov A. I.: On the possibility of using fractal models in the construction of information security systems. Computational nanotechnology 1/2017, 39–49.

[21] Ortiz S. M., Parra O., Miguel J., Espitia R.: Encryption through the use of fractals. International Journal of Mathematical Analysis 11(21)/2017, 1029–1040.

[22] Wenliang Du: Computer Security: A Hands-on Approach. CreateSpace Independent Publishing Platform, 2017.

M.Sc. Yuriy Khanas
e-mail: yuriy.y.khanas@lpnu.ua

Yuriy Khanas is an assistant at the Department of Computer-Aided Design, Lviv Polytechnic National University, Lviv, Ukraine. His main research interests include algorithms and data compression and encryption systems based on matrix transformation algebra, information protection systems and algorithms.

http://orcid.org/0000-0001-6496-5782

Ph.D. Michał Borecki
e-mail: michal.borecki@ee.pw.edu.pl

Michał Borecki received the Ph.D. degree in electrical engineering from Warsaw University of Technology, Poland, in 2017. He is an assistant professor in Department of High Voltage and Electromagnetic Compatibility.
Author of many articles in the field of lightning protection, power network and risk managemenet.

http://orcid.org/0000-0001-8907-6906

otrzymano/received: 04.03.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2413

POLYPARAMETRIC BLOCK CODING

Julia Milova, Yuri Melnyk
State University of Telecommunications, Educational-Scientific Institute of Telecommunications, Kiev, Ukraine

Abstract. *The principles of poly-parametric information coding have been considered. The methods for developing poly-parametric codes have been presented. It is shown that the protection of block codes from channel interference using check patterns can be developed by a mono- or poly-parametric method. A special type of block codes has been presented, the check patterns of which are formed on the basis of their neighbours, which are functionally related to the given code combination. Such codes have been called poly-parametric. Binary poly-parametric ring codes, the check patterns of which are designed to detect and correct channel errors, are developed using the properties of Galois fields and on the basis of the vector shift indicators of the codewords. To obtain digital poly-parametric block codes, the properties and features of the normalized natural sequence are used. It is shown that each codeword of a binary block code can be represented as a certain positive integer in the decimal number system, which is an element of the natural sequence. Its elements on an interval that equals the norm acquire a functional dependency.*

Keywords: codeword, vector shift indicators, natural sequence, poly-parametric codes

POLIPARAMETRYCZNE KODOWANIE BLOKOWE

Streszczenie. *Rozważono zasady poliparametrycznego kodowania informacji. Przedstawiono metody tworzenia kodów poliparametrycznych. Wykazano, że ochrona kodów blokowych przed zakłóceniami kanałowymi za pomocą wzorców kontrolnych może być realizowana metodą mono- lub poliparametryczną. Przedstawiono specjalny typ kodów blokowych, których wzorce kontrolne są tworzone na podstawie ich sąsiadów funkcjonalnie związanych z daną kombinacją kodową. Takie kody zostały nazwane poliparametrycznymi. Z wykorzystaniem własności pól Galois oraz na podstawie wskaźników przesunięcia wektorowego słów kodowych, zostały opracowane binarne poliparametryczne kody pierścieniowe, których schematy kontrolne przeznaczone są do wykrywania i korekcji błędów kanałowych. Do otrzymania cyfrowych poliparametrycznych kodów blokowych wykorzystuje się właściwości i cechy znormalizowanego ciągu naturalnego. Pokazano, że każde słowo kodowe binarnego kodu blokowego może być reprezentowane jako pewną dodatnią dziesiątkową liczbę całkowitą, która jest elementem ciągu naturalnego. Jego elementy w przedziale równym normie uzyskują zależność funkcyjną.*

Słowa kluczowe: słowo kodowe, wskaźniki przesunięcia wektorów, ciąg naturalny, kody poliparametryczne

Introduction

Block codes play a vital role among a large number of different methods for coding digital information. Although block codes have the most ancient history in terms of time creation, they have not lost their significance even now [7, 9]. They are mainly used for coding and exchange of book documentary information that requires special accuracy. This defines the usual structure of block codes, consisting of two parts – useful information and check symbols for error detection and correction. The symbol error correction is built on the informational part of the codeword and is often built into the general structure of the block. Over the years of its existence, truly unique, sophisticated methods of creating check patterns have been developed, which were often named after their developers – Hamming, Halley, BCH codes, etc. All of these codes use the information of only the useful part of the transmitted block [5, 10, 14]. This imposes certain limitations on the creation of the check part of a codeword separate block.

1. Fundamentals for constructing polyparametric binary block codes

Let us consider the following example. Let the useful information, presented in a block of length N binary symbols be bitwise shifted left or right N-1 times. Moreover, after each shift, the number of bitwise coinciding units in the original and offset numerical vectors is determined (V). It is clear that after such actions, we will receive the vector shift indicators (VSI) consisting of N-1 numeric elements. Note that the original vector and its N-1 close equivalents have a linear relationship with each other, which can be used to create check patterns. Individual elements of the vector shift indicators can be subjected to linear and non-linear operations, the results of which can also be used to obtain a check pattern. The considered example is shown in Fig. 1.

Fig. 1 shows that in this case, to obtain a check pattern, the structure of not only the original vector is used, but also the other N-1 shift vectors that functionally depend on it. This provides an advantage when creating a check pattern and expands the possibilities for obtaining it [4, 8, 12].

$$V = \begin{bmatrix} 1&0&0&0&0&0&1&1&1\\ 1&1&0&0&0&0&0&1&1\\ 1&1&1&0&0&0&0&0&1\\ 1&1&1&1&0&0&0&0&0\\ 0&1&1&1&1&0&0&0&0\\ 0&0&1&1&1&1&0&0&0\\ 0&0&0&1&1&1&1&0&0\\ 0&0&0&0&1&1&1&1&0\\ 0&0&0&0&0&1&1&1&1 \end{bmatrix} \quad VSI = \begin{bmatrix} 2&4&6&8&8&6&4&2\\ &2&4&6&8&8&6&4\\ &&2&4&6&8&8&6\\ &&&2&4&6&8&8\\ &&&&2&4&6&8\\ &&&&&2&4&6\\ &&&&&&2&4\\ &&&&&&&2 \end{bmatrix}$$

Fig. 1. Receiving the elements of the vector shift indicators

In other words, there are mono- and poly-parametric code combinations:

- mono-parametric block codes are the codes in which the check pattern of each block is formed only on the basis of its internal structure;
- polyparametric block codes are the codes, the check pattern of which are formed by a given block and a set of neighbouring codewords functionally depending on a given codeword.

When block codes are used, information is transmitted by codewords of constant selected length L. The check patterns of these codes are based on an internal structure, namely, the distribution of 0s and 1s within one codeword.

Let us consider the structure of 0s and 1s located between the first and last single symbols of the codeword called the delta factor. A delta factor structure of some type allows ring codes with special properties to be created. Namely, when expanding the size of the codeword by several zero symbols, the vector shift indicators are completely determined by the type of the delta factor, which enables codes with special properties to be obtained. These codes can be used as entropy codes.

The information resource for block codes is a relative value indicating the number of units of its information part, for which a check pattern is created for channel error detection and correction. For mono-parametric codes, the information resource has a limit equal to one. For polyparametric codes, it equals |L/2|+1, where L is the length of the codeword.

The information reserve is the amount of information for obtaining a check pattern in block codes, which exceeds the capabilities of the useful information part of the transmitted codeword. For mono-parametric codes, the information reserve is equal to zero, and for poly-parametric codes interconnected by functional dependency, it is equal to |L/2|-1.

Earlier, it was shown that the construction of poly-parametric codes requires a functional dependency between the received code combination and its nearest neighbours. In Fig. 1, there are N-1 of such nearest neighbours equal to the length of the code combination. Thus, binary codes that are segments of Galois fields have functional dependency [1–3].

2. Polyparametric digital decimal block codes

But it turned out that decimal codes, which are elements of the normalized natural sequence, have a similar functional dependency [9].

Deviations of the characteristics of the normalized natural series from the non-normalized ones can be used as parameters of the set of codewords built on their basis.

Such features of the normalized natural series create conditions for varying its elements in order to create new codes with new characteristics. These characteristics can be used as parameters.

Polyparametric codes are closely related to the usual natural series, which is an infinite sequence of integers differing by one on a unit interval.

Under the concept of a normalized natural sequence, a natural sequence should be taken after each of its elements by the same integer. It is convenient to choose this integer equal to the length of the information part of the transmitted block. After normalization, the natural sequence begins to possess the following features:

- elements of the natural sequence are converted into real numbers, consisting of an integer part (modulus) and a fractional part (remainder);
- with respect to the remainders, the normalized sequence or its segments are divided into sections of the same length equal to the size of the norm, which we have proposed calling cycles;
- in all cycles of the normalized natural sequence and in its segments, the remainders are repeated and depend only on the size of the norm;
- cycles are separated from each other by a normalized element with a zero remainder.

When an element of the natural series becomes equal to or a multiple of the value of the norm, its remainder becomes zero, and the value of the integer part is increased by one and becomes equal to the ordinal number of the next cycle. Up to this point, the integer part of all normalized elements is equal to the cycle ordinal number reduced by one. The fractional part of each normalized element in the cycle is determined only by the value of the norm.

Below are examples of two cycles of normalized natural series for two different norms and two different lengths of clippings.

These features are shown in Fig. 2.

Here, Gx(n) and Gx(k) are the cycles obtained by normalizing the natural sequence, L = 7 and 11 are the size of the norm, and *n* and *k* are the lengths of the line segment of the normalized natural series, from which cycles are selected. The parameters of the given cycles are as follows:

G(n), L = 7, n = 70…84
G1(n), L = 11, n = 70…84
G(k), L = 7, k = 56…70
G1(k), L = 11, k = 56…7

If you select elements of the normalized natural sequence as decimal codes, they will have the abovementioned properties. It can be considered that any deviation of the properties of the normalized natural sequence and various combinations of its elements are the basis for getting new parameters of codewords obtained due to this sequence or its normalized equivalent. Such parameters, with the help of adjacent codewords, enable check patterns to be obtained for channel error detection or correction.

In particular, the elements of the normalized natural sequence, used as code combinations, receive two parameters: the ordinal number of the cycle from the beginning of the natural sequence in which the code combination is located and which can be considered as the index of this combination, as well as its number from the beginning of the cycle, used as an offset. Using these parameters, it is easy to check the correctness of the code combination and to correct existing errors.

It is worth noting the fundamental feature of the given parameters: their one-to-one correspondence with the parental code combination. In fact, these two parameters can be transmitted with equal success instead of the parent code combination, and vice versa.

$n = 70..84$		$k = 56..70$	
$G(n) = \frac{n}{7}$	$G1(n) = \frac{n}{11}$	$G(k) = \frac{k}{7}$	$G1(k) = \frac{k}{11}$
10	6.364	8	5.091
10.143	6.455	8.143	5.182
10.286	6.545	8.286	5.273
10.429	6.636	8.429	5.364
10.571	6.727	8.571	5.455
10.714	6.818	8.714	5.545
10.857	6.909	8.857	5.636
11	7	9	5.727
11.143	7.091	9.143	5.818
11.286	7.182	9.286	5.909
11.429	7.273	9.429	6
11.571	7.364	9.571	6.091
11.714	7.455	9.714	6.182
11.857	7.545	9.857	6.273
12	7.636	10	6.364
…	…	…	…

Fig. 2. Cycles of a normalized natural sequence

3. Coefficients before normed codes

Problems of coefficients in front of normalized codes arise due to the fact that the coefficient "k" located in front of the value of an element of the natural series is divided by the norm L. As a result of this, at least three varieties of the quotient from the division are formed: the quotient is equal to, less than or greater than one. Moreover, the value of the quotient can be very different in magnitude.

The first kind of quotient occurs when the norm is the largest multiple of the coefficient or is equal to it. For example:

*84*k:7=12*k*
*48*k:8=6*k*
*17*k:17=k*

The second type of quotient occurs when a quotient is a fractional number greater than or less than one. For example:

*67*k:13=5.154*k*
*13*k:24=0.542*k*

The third kind of quotient appears when the numerator and denominator have a common factor. Then the quotient becomes basically a real number. Its peculiarities, in this case, are preserved. For example:

*66*k:26=33*K:13*
*9*K:39=3*K:13*

Four examples of the resulting cycles of a normalized natural series with coefficients in front of the elements are given below:

The sampling interval of the natural range is from 65 to 85 elements.

Figure 3 shows the cycles in more detail.

$$n := 67..85$$

$$G5(n) = \frac{3n}{8} \quad G6(n) = \frac{3n}{10} \quad G2(n) = \frac{3n}{11} \quad G8(k) = \frac{3n}{13}$$

$G5(n) =$	$G5(n) =$	$G5(n) =$	$G5(n) =$
25.125	20.1	18.273	15.462
25.5	20.4	18.545	15.692
25.875	20.7	18.818	15.923
26.25	21	19.091	16.154
26.625	21.3	19.364	16.385
27	21.6	19.636	16.615
27.375	21.9	19.909	16.846
27.75	22.2	20.182	17.077
28.125	22.5	20.455	17.308
28.5	22.8	20.727	17.538
28.875	23.1	21	17.769
29.25	23.4	21.273	18
29.625	23.7	21.545	18.231
30	24	21.818	18.462
30.375	24.3	22.091	18.692
...	...	...	...

Fig. 3. Coefficients of elements of normalized natural series

From the examples above, you can see that:

- the cyclicity of the normalized natural series remains, which is equal to the size of the norm;
- each time, the product "coefficient multiplied by the current value of the row element" is subjected to normalization;
- a normalized natural series with coefficients relative to the fractional part (remainders) is divided into segments of the same length – cycles separated from each other by elements with zero remainder;
- the residuals during normalization of the natural series with coefficients are formed not for each successive element of the natural series, but for those elements that are created after multiplying each element by the corresponding coefficient;
- cycles are equal to the value of the norm;
- if the coefficients are real numbers, then the normalized natural series loses its cyclicity;
- a normalized natural series with coefficients at a rate of 10 forms normalized elements with residuals of one digit (multiple of 10).

4. Constructing polyparametric codes

Considering the line segments of the normalized natural series with coefficients in front of each element, we come to the conclusion that, by combining the elements of the natural series and then normalizing them, it is possible to construct new polyparametric digital codes with two or more parameters. An example of the result of such a construction is the so-called summary code and its varieties. Each code combination of the total code is obtained by arithmetic addition of all elements of the natural series up to this number, including it. It is clear that if the n-th code combination is created (constructed), it will be equal to the sum 1 + 2 + 3 + … + n. The sequence of code combinations obtained in this way is then normalized using the selected rate.

For the summary codes, the following patterns are observed:

- the normalized natural series, as well as its individual segments with respect to the fractional part (remainders) of neighbouring elements, splits into cycles-segments of the same length, equal to the size of the norm L;
- within each cycle with respect to its middle part, for all remainders, the even symmetry is established: at the same distance from the centre of the cycle, the residuals of normalized elements are equal to each other;
- in each cycle, there is a pair of normed elements with zero residuals, which can be used to frame the cycles;
- if the norm is an even number, then, relative to the remainders, each element of the cycle has its own pair, with an odd norm, the middle element of the pair does not have.

These patterns can be seen in more detail in Fig. 4.

The sequences designated S3, S4 and S7 represent cycles of normed codewords at rates equal to 6, 8, and 9, respectively.

Here, the cyclic essence of the normed natural series is immediately revealed, and the cycle size is equal to the norm, as well as zero residuals of codes at the ends of the cycles (however, the beginning of the cycles can be selected from any element in the cycle and this will be correct). It is more convenient to start and end cycles with zero remainders. One asymmetric element is well traced with remainders 0.857 for the L = 7 norm and 0.222 for the L = 9 norm. For an even norm, such as L = 8, the remainders of all normed elements in the cycle are repeated twice.

S3(n) L=70 0; 0.143; 0.429; 0.857; 0.429; 0.143; 0
S4(n) L=8 0.25; 0.75; 0.5; 0.5; 0.75; 0.25; 0
S7(n) L=9 0; 0.222; 0.667; 0.333; 0.222; 0.333; 0.667; 0.222; 0.

$$n := 35..54$$

$$S3(n) = \sum_{k=1}^{n} \frac{k}{7} \quad S3(n) = \sum_{k=1}^{n} \frac{(2k)}{8} \quad S3(n) = \sum_{k=1}^{n} \frac{(2k)}{9}$$

$G3(n) =$	$G4(n) =$	$G7(n) =$
90	157.5	140
95.143	166.5	148
100.429	175.75	156.222
105.857	185.25	164.667
111.429	195	173.333
117.143	205	182.222
123	215.25	191.333
129	225.75	200.667
135.143	236.5	210.222
141.429	247.5	220
147.857	258.75	230
154.429	270.25	240.222
161.143	282	250.667
168	294	261.333
175	306.25	272.222
...	...	

Fig. 4. Total code combination loops

5. Digital total codes and their features

It is known that any binary code combination can be represented by a positive integer. Since the total codes are numerical, and in their structure, each total codeword is derived from a natural number series, we agree to number the counting set of total codes in order with the elements of the natural series 1, 2, 3, …, n and form them by simply adding the elements of this series. Also, the total codeword can be easily represented as a binary number [13]. For example, a code with sequence number n = 5 is obtained as a binary representation of the total of decimal numbers S(n) = 1 + 2 + 3 + 4 + 5 = 15, i.e. as Z = 1111, and a code with sequence number n = 10 is encoded as a binary representation of the total of decimal numbers S(n) = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55, namely Z=110111.

A feature of total codes is that they are polyparametric. Each codeword typically has several parameters by which it can be extracted from a plurality of similar codewords and, against the background of external interference distorting the codeword structure, reconstructed. Let us indicate two main parameters of total codes, which make it possible to check the correctness of the code word of the total code and in many cases restore it to its original form:

- belonging of the codeword to the group of total codes. If the received codeword, in accordance with its value, does not belong to the group of total codes constructed in accordance with

the above algorithm, it is erroneous. The sum structure of the codeword is its main parameter;
- the remainder of dividing the total codeword by a given number. If the remainder of dividing the codeword by some selected pre-known integer has a remainder other than the expected remainder, it is erroneous. The type and value of the remainder of dividing the total codeword by a given number is its second parameter.

There are two properties of total codes that can be easily verified experimentally.

1) A set of codewords on a certain interval of their lengths n, being correlated to any small integer K, necessarily gives one or several dual multiplicities. Dual multiplicities D are the pairwise results of dividing the total codeword with the number of elements n by K without the remainder. Dual multiplicities are quite common. Therefore, such code words are convenient to use in practice.
2) If one integer is divided by the second, the result is the whole part and the remainder, which is always less than the divisor. For total codes, there is such a pattern that, starting from any dual multiplicity, the remainders of dividing integers of the total codes S (n) by the divisor K up and down the set of ordinal numbers n are symmetric and pairwise equal to each other, starting from any dual multiplicity [11].

The remainder values are stored over the entire set of codeword numbers.

Thus, for each total codeword, there are three of its identification parameters:
- serial number n;
- value of the total codeword S(n);
- remainder of dividing the total codeword by the selected number k.

Note that two adjacent dual-fold total codewords have zero residuals.

There are three types of total codes: natural numbers, even and odd total codes. The last two codes can be considered as derivatives of the total natural number code.

Even total codes are formed on the basis of even numbers of natural numbers 1, 2, 3, … n according to the formula 2*n, and odd total codes according to the formula 2*n + 1. Here, the next codeword is recursively obtained from the previous codeword by adding the next even or odd number, respectively. The result is a sequence of codewords. In particular, the first ten codewords of an even total code have the form 2 6 12 20 30 42 56 72 90 110. Similarly, the first twelve odd codewords form a sequence of 3 7 13 21 43 57 73 91 111. As follows from the logic of things, code combinations with the same sequence numbers differ by one, and the odd summary codes prevail over the even ones. Therefore, the basic properties in both codes must be the same.

The properties of even and odd sum codes demonstrate the following patterns.
- at any interval, the normalized sequence of the sum code breaks down into cycles of equal length, the size of which is determined by the size of the norm; each cycle is bordered by an integer multiple of the norm value;
- normalized code words in cycles are real numbers, the fractional part of which is a set of values symmetric with respect to the average code and equal to each other.

These fractional parts for each norm-different and are stored as constants for all received cycles on an infinite sequence of codewords.

6. Conclusions

The protection of block codes from channel interference using check correcting patterns can be performed in a mono-parametric or poly-parametric way, i.e. either one code combination at a time or using the code combinations of its neighbours functionally related to it.

For this purpose, a methodology for creating binary and digital decimal poly-parametric codes have been proposed.

Binary poly-parametric codes are developed using the properties of Galois fields, whereas digital block codes are based on the natural sequence. In this case, the basis of poly-parametric codes (Galois fields and the original natural sequence) is subjected to linear transformations to obtain a functional dependency of the neighbouring codes.

Based on the thinned natural sequence, poly-parametric codes can be developed to detect, correct errors, and protect from unauthorized access.

For detecting and correcting combinations of poly-parametric codes, new possibilities open up due to additional information embedded in adjacent codewords.

Poly-parametric codes, in comparison to mono-parametric codes, using only one code combination, simplify, improve and diversify the choice of code protection from errors and interference.

References

[1] Arora S., Barak B.: Computational Complexity: A Modern Approach. Cambridge University Press, Cambridge 2009.
[2] Berlekamp E.: Algebraic coding theory. Mir, Moscow 1971.
[3] Blahut R. E.: Algebraic Codes for Data Transmission. Cambridge University Press, 2012.
[4] Bronshtein I. N., Semendyaev K. A.: Mathematics reference book for engineers and students of technical colleges. GITTL, Moscow 1957.
[5] Brouwer E., Shearer J. B., Sloane N. J. A., Smith W. D.: A new table of constant weight codes. IEEE Trans. Inform. Theory 36/1990, 1334–1380.
[6] Carrasco R. A., Johnston M.: Non-binary error control coding for wireless communication and data storage. J. Wiley & Sons, 2008.
[7] Conway J. H., Sloane N. J. A.: Lexicographic codes: error-correcting codes from game theory. IEEE Trans. Inform. Theory 32/1986, 337–348.
[8] Dikarev A. V.: Codes based on binary rings. Control systems, navigation and communication 1(29)/2014. 50–53.
[9] Etzion T.: Optimal constant weight codes over Zk. and generalized designs. Discrete Math. 169/1997, 55–82.
[10] MacKay D., Neal R.: Near Shannon limit performance of low density parity check codes. IEEE Electronics Letters 32(18)/1996, 1645–1646.
[11] Milova J. A.: Parameters of total codes. Zvyazok 4/2018, 3–32.
[12] Milova Y.: Rationed natural row. Polyparametric coding. The European Journal of Technical and Natural Sciences 3/2020, 19–23 [http://doi.org/10.5604/20830157.1121333].
[13] Milova J. A. et al.: Total codes. Zvyazok 3/2018, 47–50.
[14] Robinson J. P., Bernstein A. J.: A class of binary recurrent codes with limited error propagation. EEE Transactions on Information Theory 13(1)/1967, 106–113 [http://doi.org/10.1109/TIT.1967.1053951].

M.Sc. Julia Milova
e-mail: milovajul25@gmail.com

Ph.D. student of the Educational-Scientific Institute of Telecommunications of the State University of Telecommunications. Teaches at the Kiev College of Communications. Works on the development of information coding methods.

http://orcid.org/0000-0002-9977-2195

Ph.D. Yuri Melnyk
e-mail: milovajul25@gmail.com

Doctor of Technical Sciences. Director of the Educational-Scientific Institute of Telecommunications of the State University of Telecommunications. Scientific direction – telecommunications network management. Deals with issues of hierarchical management of telecommunications networks and issues of fuzzy logic in management problems. Development of the human-operator information model in ergatic systems.

http://orcid.org/0000-0002-5028-8749

otrzymano/received: 30.11.2020 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2429

NO-CODE APPLICATION DEVELOPMENT ON THE EXAMPLE OF LOGOTEC APP STUDIO PLATFORM

Monika Moskal
Silesian University in Katowice, Institute of Sociology, Katowice, Poland

Abstract. *Digitalisation is one of the major trends which changes society and businesses. Digitalisation include buying software helping in the automatization of business process. Gaining competitive advantage can be achieved by implementing dedicated (customized) IT solutions tailored to the specificity of a given enterprise. However, dedicated solutions are expensive, and many companies cannot afford them. No-code technology – the newest trend in the IT sector – are said to be the solution to this problem. The article discusses why no-code technology is gaining popularity. The paper also presents the differences between no-code and traditional programming approaches. A detailed description of no-code technology was based on the investigation of a no-code platform called Logotec App Studio. The article stresses the possible applications of no-code tools in companies' digitalisation process.*

Keywords: no-code technology, no-code application development platform, Logotec App Studio

TWORZENIE OPROGRAMOWANIA BEZ KODOWANIA NA PRZYKŁADZIE PLATFORMY LOGOTEC APP STUDIO

Streszczenie. *Cyfryzacja to jeden z głównych trendów zmieniających społeczeństwo i funkcjonowanie przedsiębiorstw. Cyfryzacja obejmuje zakup oprogramowania pomagającego w automatyzacji procesów biznesowych. Uzyskanie przewagi konkurencyjnej jest możliwe poprzez wdrażanie dedykowanych (niestandardowych) rozwiązań informatycznych dostosowanych do specyfiki danego przedsiębiorstwa, jednakże ze względu na wysokie koszty pozostają one poza zasięgiem większości przedsiębiorstw. Technologia no-code – najnowszy trend w branży IT – jest rozwiązaniem tego problemu. W artykule zostały omówione przyczyny, dla których technologia no-code zyskuje na popularności. Przedstawiono również różnice między tradycyjnym podejściem do tworzenia rozwiązań IT, a podejściem no-code. Szczegółowy opis technologii no-code został oparty na badaniu platformy no-code o nazwie Logotec App Studio. W artykule podkreślono możliwości zastosowania narzędzi no-code w procesie cyfryzacji przedsiębiorstw.*

Słowa kluczowe: technologia no-code, platforma do tworzenia aplikacji bez kodowania, Logotec App Studio

Introduction

In today's world, large amounts of data are processed and stored every day [7]. Digitalization is one of the major trends which changes society and businesses. The process of digitalization causes changes for businesses due to the adoption of digital technologies in the organization or in its' environment [15]. According to literature, the process digital transformation refers to "the changes associated with the application of digital technology in all aspects of human society" [17]. According to Keiss and Brennen [3], digitalization is "the adoption or increase in the use of digital or computer technology by an organization, industry, country, etc.". Digitalization may include turning existing products or services into their digital variants [8] or buying software helping in the automatization of business process. Gaining competitive advantage can be achieved by implementing dedicated IT solutions tailored to the specificity of a given enterprise. However, dedicated solutions are expensive, and many companies cannot afford them. No-code and low-cod application development platforms are said to be the solution to this problem. Both are said to be disruptive technologies which will change the process of application development. For instance, according to Gartner, the leading technology research and advisory company [6] "by 2024, three-quarters of large enterprises will be using at least four low-code development tools for both IT application development and citizen development initiatives" [18]. The current study investigates the possibilities, advantages and potential disadvantages of no-code technology based on the example of Logotec App Studio – no-code application development platform.

1. No-code – the future of application development

1.1. No-code and low-code programming differences

In the beginning, it is important to make a distinction between low-code and no-code development platforms, which can seem unclear due to the varying nature of platforms' functionalities [2] and the terminology which might be misleading. Low-code programming has become the fourth generation of programming languages. It is also called declarative programming. In this programming approach, it is possible to create IT solutions without extensive programming knowledge when comparing it with the traditional programming approach [1]. However, programming language knowledge is necessary to use low-code development platforms. In contrast, no-code approach does not require any programming knowledge, but the ability of logical and abstract thinking is necessary to create software solutions without coding. The core design of the software solutions in no-code approach is made by the developer through drag and drop manipulation. However, low-code platforms make the creator to be more dependent on the hard core for dictating the application's core architecture. [5] There are many no-code and low-code tools on the market which differ widely in their functionality. However, there is a clear trend: there is a disappearing distinction between tools simple enough for citizen developers and adequately powerful for professional development teams. Therefore, no-code and low-code will combine into a single marketplace segment [2].

1.2. No-code platforms – the solution for a software crisis

The introduction to almost every text on software development or software engineering starts with a Software Crisis description. In general, the criticisms are that software projects are over budget, incorrect, late and do not do what they intended [4, 9, 10]. According to Molyneux [13], similar problems arise in order, broader areas of IT or implementing and designing software systems. The reasons why no-code application development platforms gain popularity can be divided into four main aspects: the scarcity of qualified programmers, continuously changing technologies and the need to learn them, high costs, time-consuming and complicated process of software development. The first reason why no-code is gaining popularity is the scarcity of qualified programmers which hinders the rapid implementation of the project. For a long time, it has been noticed that a large number of students find their (first, second, and third) contact with computing and programming daunting. According to Knuth [10], many programming teachers have observed that only a very small percentage of students really "resonate" with the subject. Finding a qualified (or high-class) programmer does not guarantee

a project's success. The programmers and software engineering while producing quality software often encounter other problems, for example, how to produce the software within the budget, which is a difficult task [13]. No-code tools, on the other hand, as they do not require programming knowledge, can be used by anyone capable of algorithmic thinking. Therefore, no-code platforms can not only be used by programmers to speed up the process of software development but also by citizen developers, business analysts, consultants, or solutions designers to create business applications without a line of code, resolving the problem of scarcity of qualified programmers on the market. Second, continuously changing technologies and the need to learn them is another factor which causes the rise of no-code popularity. Developers learn a chosen programming language which, apart from long years of studying, often requires many years of practice before one can create advanced IT software. A technological switch also requires a long learning process and many hours of practice which for some may be discouraging. Therefore, only a few of qualified programmers can switch to another programming technology. Third, developing IT solutions using traditional approaches requires large financial outlays. Even system engineers or software engineers notice that developing quality software within set budget is a very difficult activity [13]. However, no-code offers solutions at an incomparably lower cost which will be discussed later in the article. Moreover, programming in traditional methods is also very time-consuming. The time needed to create IT software (for one operating system and device) using traditional programming is often not shorter than 6 moths and very often the time needed to create solutions running under all popular operating systems requires years of work of a team of professionals. The no-code approach allows to develop software solutions easily and quickly without the need to wait for IT professionals to complete all programming services needed [1]. It is important to note that different no-code platforms offer different functionalities which for one project the functionalities can be disadvantages and for others – very useful features. However, to make an in-depth analysis, one no-code platform (Logotec App Studio) was chosen for further investigation. Around a year and a half was spent to study the platform and its functionality. Moreover, the author used the platform to develop applications with varying functionalities.

2. No-code Development Platform on the example of Logotec App Studio

The subject of a study was Logotec App Studio which is a no-code development platform created by Logotec Group. The company is an application generator expert – it has been doing it for 32 years (the company received its first awards for the "Report and Application Generator" in 1986 and 1989. At the beginning of the 21st century, Mobile@Connector was created. It was an IT tool enabling creating IT software for mobile devices (Pocket PC at that time) completely without programming. Logotec Group with its Mobile@Connector received a number of international awards, including an award from Bill Gates in 2003 in New Orleans for the best tool for creating mobile applications [19]. Unfortunately, in 2010, Microsoft has stopped the development of Pocket PC and despite many awards, the company in order to survive, had to make a technological switch. In the last few years, Logotec Group has created – using new technologies – a new IT solution – Logotec App Studio.

Unlike Mobile@Connector, Logotec App Studio enables the creation of solutions not only for mobile devices (smartphones), but for all devices (computers, tablets, smartphones) and for all popular operating systems, i.e.: Windows 10, macOS, Android, and iOS. It is a no-code platform which enables the development of "azure" database (business) applications without a line of code. With this platform, IT software is created in two orders of magnitude shorter time than with traditional methods and two orders of magnitude cheaper than with traditional methods. The platform's creators use a data first approach which means that an application is generated based on previously created tables in a relational database. Theoretically, a database can be any data source. However, currently the platform allows to choose Microsoft SQL Server, Oracle, MySQL, PostgreSQL, MariaDB. The server layer is a .NET Core, Windows or Linux (Microsoft Azure) however, the client layer runs on Windows 10 (Universal Windows Platform). Logotec App Studio is multilingual (currently available languages: English, German, Polish, and Chinese) and was prepared for translation into any other language. The application is defined once (without programming) and then, the platform (residing in the cloud) automatically generates ready-to-use, fully functional, multilayer "azure" applications running all popular operating systems. The generated solution is a native not a browser application. Created applications work online as well as offline. When the Internet connection is lost, the application turns on an offline mode which allows the user to continue the work. However, when the connection is restored, all data are synchronized and uploaded to server.

2.1. Logotec App Studio versus traditional programming approach

Differences between traditional programming and no-code platform (Logotec App Studio) in eight categories are presented in Figure 1. The size of implementation team when using traditional programming methods requires a team of professionals. However, when no-code is used, only one person is required to create IT software.

	Traditional approach	Logotec App Studio
Implementation team	**A team** of professionals	**One** person
Required competencies	**Extensive** programming experience (Windows 10, Android, iOS, macOS)	Programming skills **not required** (e.g. business analyst)
Development time (prototype)	2-3 **months**	2-5 **days**
Cost (prototype)	>> **200** man-days	**2-5** man-days
Development time (final solution) (Windows 10, Android, iOS, macOS)	A few **months**	A few **days**
Cost (final solution) (Windows 10, Android, iOS, macOS)	>> **1000** man-days	**A few** man-days/man-weeks
Modification	**Difficult** **Time-consuming** **Expensive**	**Minutes-hours** Designed for easy modifications
Investment risk	**Very high** (among others because of the cost of the prototype)	**Minimal** (prototypes generated ad-hoc)

Fig. 1. The differences between traditional programming approach and Logotec App Studio (no-code platform) [11]

Moreover, a team of professionals (which uses traditional programming methods) has to have an extensive programming experience for all operating systems, which is not a requirement in the case of Logotec App Studio because an application defined once works automatically on all popular operating systems (Windows 10, Android, iOS, macOS) and can be easily (without coding) modified. The time needed to develop IT solutions with no-code platform shortens from months to days.

The time needed to prepare the prototype ranges from hours to a few days. The cost of creating both the prototype and the final solution is incomparably lower than using traditional programming methods. For instance, more than 200 man-days are required to create a prototype and more than 1000 man-days are needed to create a final solution. However, when using no-code platform, only 2–5 days (in practice, sometimes only a few hours) are needed to create a minimum viable product (MVP) and a few man-days or man-weeks are needed to create a final solution. Moreover, modifications of a product developed with traditional methods are very time-consuming, expensive and difficult. However, in the case of disruptive technology, the modifications are quick and easy. The final category is investment risk which is very high when using traditional programming methods. However, there is low investment risk when discussing no-code – some no-code platform authors offer free MVP's (which are generated ad hoc) or the possibility of resigning from using the application at any time without incurring further costs. Moreover, some platforms, like Logotec App Studio, are offered in the SaaS (Software as a Service) model making investment risks even lower.

2.2. Logotec App Studio versus traditional programming

Finances are important factors influencing decisions on the process of digitalization. Only a small proportion of companies are able to buy dedicated software. Nevertheless, the disruption of no-code technology is that it changes the situation. Thanks to an ability to develop software without a line of code, the process becomes cheaper and less time-consuming. This opens the door to small and medium companies. Figure 2 represents the changes in cost per user of a standard solution depending on the size of the software project. Costs per user with the traditional approach grow exponentially as the number of users decreases. However, it does not concern projects implemented with no-code platforms – costs per user are incomparably low.

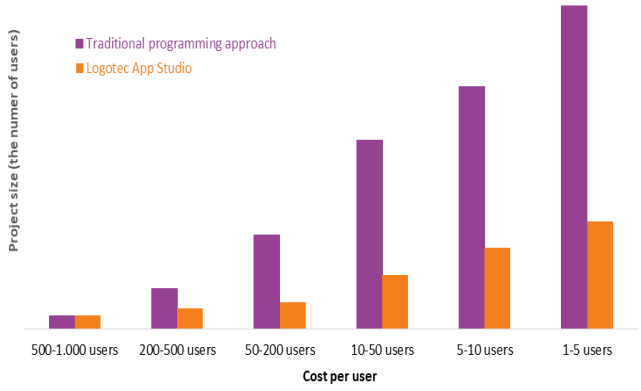


Fig. 2. Costs per user of a typical solution depending on the size of the project

The targeted software project size of traditional programming is different from the project size no-code platforms are aimed for. The traditional programming approach becomes unattractive in terms of price as the number of potential users of the solution decreases. It is attractive only when the number of users is high. However, the no-code gains the highest attractiveness in small software projects with small number of users (see Figure 3).

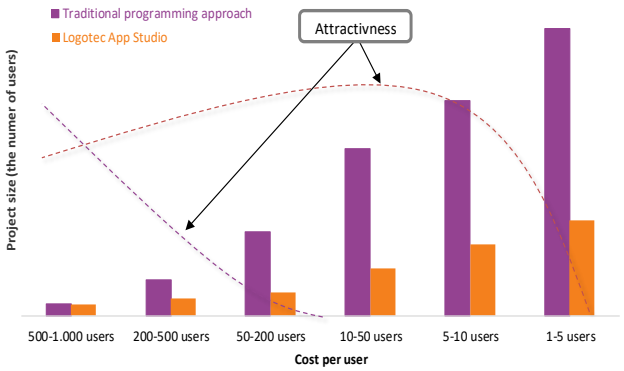


Fig. 3. The value of the solution depending on the size of the project

Small or medium software projects include dedicated software for small or medium sized companies or a specific application for companies' departments. The high costs of dedicated software allow only big companies (with more than 200 persons employed) to afford it, leaving small and medium companies without a possibility to digitalize. Nevertheless, these companies (and projects) constitute a much greater proportion of the marketplace than large enterprises employing more than 200 persons) (see Figure 4) [14]. A disruptive no-code technology resolves this problem. By providing a tool with which almost anyone can develop a native business application for all popular operating systems and devices the door to digitalization is open to those to whom it was unobtainable (due to many reasons mentioned above) before.

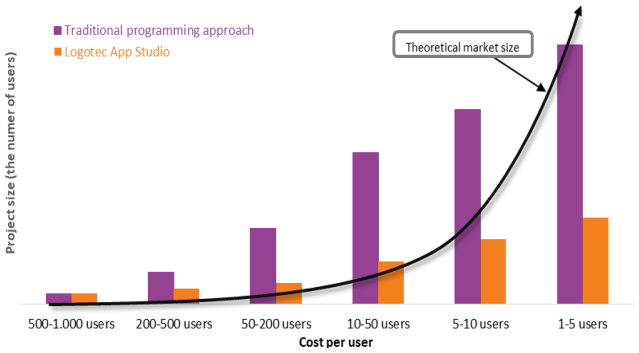


Fig. 4. Potential / theoretical size of the market

3. No-code as a disruptive technology

No-code and low-code are disruptive technologies gaining rising popularity. Using no-code platform, like Logotec App Studio, a person without any programming knowledge is able to develop IT software for all devices and all popular operating systems. However, professional programmers can also benefit from no-code technology. The platforms allow qualified programmers to broaden their portfolio by being able to create more projects in a shorter time without spending many months to develop a software. The costs of developing an application using no-code is cheaper and quicker than when using the traditional programming approach. There is low investment risk – some no-code platform authors, like Logotec Group, offer free MVP's or the possibility of resigning from using the application at any time without incurring further costs. Moreover, some platforms, like Logotec App Studio, are offered in the SaaS (Software as a Service) model making investment risks even lower. The changes in software developed with no-code platforms are implemented ad hoc and they can be made by anyone knowing the platform without programming knowledge. No-code platforms differ in their functionality. Some offer the development of browser solutions, the other – native applications; some require the data-first approach, preparation

of the tables of relational database based on which an application is generated, and the others – do not require the database preparation; some platforms offer generating applications for mobile devices, while the others – for all devices and most of operating systems; some may rise data privacy concerns, while the others employs multiple data protection mechanisms, etc. Therefore, when looking for a no-code tool it is important to analyse and compare it with others' tool functionalities and features in order to choose the one which would meet one's needs.

4. Conclusions

The world's situation with coronavirus pandemics, contrary to the appearance, seems to be a significant opportunity for no-code technology which makes dedicated software available to the most of the companies present on the market. The pandemic forces a great proportion of companies to radically lower or stop investment. Becoming even more cautious with one's decisions, increases the awareness of the labour costs and the risks associated with it. This leads to an increase in the readiness to digitalize and automate work, products and services. This digitalization in the form of introducing dedicated software gives many possibilities, for instance: enables employees to work without leaving home, prevents data loss when employees are in quarantine, etc.

The no-code technology is definitely disruptive, and soon we will see, if the Gartner's statement that "by 2024, three-quarters of large enterprises will be using at least four low-code development tools for both IT application development and citizen development initiatives" [18] will be true.

M.Sc. Monika Moskal
e-mail: moni.moskal@outlook.com

A Ph.D. candidate at Institute of Sociology at Silesian University in Katowice and IT Business Analyst using no-code platform in creating native business database software solutions. Research interest include IT (especially no-code technology) and assimilation processes.

http://orcid.org/0000-0002-4640-6095

otrzymano/received: 11.12.2020 *przyjęto do druku/accepted: 15.03.2021*

Bibliography

[1] Adrian B., Hinrichsen S., Nikolenko A.: App Development via Low-Code Programming as Part of Modern Industrial Engineering Education. Advances in Human Factors and Systems Interaction: Proceedings of the AHFE 2020 Virtual Conference on Human Factors and Systems Interaction. Springer Nature 2020, 45–54.
[2] Bloomber J.: The Low-Code/No-Code Movement: More Disruptive Than You Realize. Forbes [https://www.forbes.com/sites/jasonbloomberg/2017/07/20/the-low-codeno-code-movement-more-disruptive-than-you-realize/#52be2098722a] (20 July 2017).
[3] Brennen S., Kreiss D.: Digitalization and Digitization. [http://culturedigitally.org/2014/09/digitalization-and-digitization/] (8 September 2014).
[4] Buxton J., Macro A.: The Craft of Software Engineering. Addison-Wesley, Wokingham 1987.
[5] Ciot T.: What is a Low-Code/No-Code Platform? [https://devops.cioreview.com/cxoinsight/what-is-a-lowcodenocode-platform-nid-15249-cid-99.html] (17 November 2020).
[6] Gartner Inc.: Gartner. [https://www.gartner.com/en/about].
[7] Han J., Pei J., Kamber M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers Inc., Burlington 2011.
[8] Henriette E., Mondher F., Boughzala I.: The Shape of Digital Transformation: A Systematic Literature Review. Ninth Mediterranean Conference on Information Systems (MCIS). Samos 2015.
[9] Ince D. C.: Software Engineering. Van Nostrand Reinhold, London 1989.
[10] Knuth D. E.: Algorithmic Thinking and Mathematical Thinking. The American Mathematical Monthly 92(3), 1985, 170–181.
[11] Logotec Engineering S.A. [https://logotec.pl/].
[12] Logotec Engineering S.A. [https://logotec.pl/pl/o-nas].
[13] Molyneux P.: Declarative Programming. Managing with Information Technology. Springer, London 1993.
[14] OECD: Enterprises by business size. 2021 [http://doi.org/10.1787/31d5eeaf-en].
[15] Parviainen P., Tihinen M., Kääriäinen J., Teppola S.: Tackling the digitalization challenge: how to benefit from digitalization in practice. International Journal of Information Systems and Project Management V(1), 2017, 63–77.
[16] Sommerville I.: Software Engineering. Addison-Wesley, Wokingham 1989.
[17] Stolterman E., Fors A. C.: Information Technology and the Good Life. Information Systems Research: Relevant Theory and Informed Practice. Kluwer Academic Publishers, London 2004.
[18] Vincent P., Iijima K., Driver M., Wong J., Natis Y.: Gartner Magic Quadrant for Enterprise Low-Code Application Platforms. [https://www.gartner.com/en/documents/3956079/magic-quadrant-for-enterprise-low-code-application-platf].

http://doi.org/10.35784/iapgos.2562

THE TRAINING APPLICATION BASED ON VR INTERACTION SCENARIOS – WITH EXAMPLES FOR LOGISTICS

Wojciech Włodyka, Dariusz Bober
University of Rzeszów, College of Natural Sciences, Department of Computer Science, Rzeszów, Poland

Abstract. *The main goal of the article is to share some experimental view onto VR training applications, that allows for adaptation of the directed training scenario to a user being trained. The resulted application allows to the three scenarios being selected in any combination of ones, as well as the tasks modification through the input files, before a test starts. The program was developed in the Unity engine, using the modern UnityXR framework providing extensive support for leading virtual reality hardware. The resulting solution presents the possibility of an efficient training in Virtual Reality while modifying the course of training without any needs to recompile the program, and also it shows some positive values of using VR technology as a didactic solution.*

Keywords: virtual reality, industrial training, learning systems, logistics

APLIKACJA SZKOLENIOWA OPARTA NA SCENARIUSZACH INTERAKCJI VR – NA PRZYKŁADACH DLA LOGISTYKI

Streszczenie. *Głównym celem artykułu jest podzielenie się eksperymentalnym spojrzeniem na aplikacje treningowe VR, które pozwalają na dostosowanie scenariusza szkolenia ukierunkowanego do szkolonego użytkownika. Powstała aplikacja pozwala na wybór trzech scenariuszy w dowolnej kombinacji, a także modyfikację zadań poprzez pliki wejściowe, przed rozpoczęciem testu. Program został opracowany w silniku Unity, z wykorzystaniem nowoczesnego frameworka UnityXR zapewniającego szerokie wsparcie dla wiodącego sprzętu wirtualnej rzeczywistości. Uzyskane rozwiązanie przedstawia możliwość efektywnego szkolenia z Wirtualnej Rzeczywistości przy jednoczesnej modyfikacji przebiegu szkolenia bez konieczności rekompilacji programu, a także pokazuje pozytywne walory wykorzystania technologii VR jako rozwiązania dydaktycznego*.

Słowa kluczowe: rzeczywistość wirtualna, szkolenia przemysłowe, systemy edukacyjne, logistyka

Introduction

Along with the development of augmented and virtual reality technologies, the scope of such applications is increasing significantly. Over the last two decades, many times, scientists have considered the use of this technology, for example in the fields of medicine (neurosurgery and anatomy) [7] and architecture [10].

However, only with the popularization and wider availability of VR's equipment, these solutions begin to penetrate into many industries, and also they are used in a wider spectrum of fields. We come to the point where the use of such VR solutions is even sometimes cheaper and much more attractive, than classic methods based on natural interaction with physical objects, especially in a dangerous environments.

Virtual Reality is no longer a solution for a narrow group of recipients, now it has become a competitive tool for teaching, entertainment and commercial purposes.

The authors of this article have pointed that many of the observed VR applications of training present one, predetermined course scenario. The authors formulate a hypothesis that, it is justified and possible to develop multi-variant applications, in which ones the interaction scenario will be adjusted individually to the needs of a given actor.

The assumption is, by using the indicated tools it is possible to develop a multi-variant training application in VR environment. The selected area for developing application will be a concern on a training course of an employee applying to work in a logistics warehouse. The three flexible scenarios will be implemented as examples.

1. Evolution of VR technology

The concept of Virtual Reality is not a new idea – the beginning of a commercial application of this technology is dated to the 90s of the last century [4], when the first commercial devices of this type has appeared. At the beginning, those applications were used only as entertainment devices, but even then it had been noticed the potential of this medium in other areas. Virtual Reality offers something that other medium can't guarantee – a transfer directly into the middle of a displayed environment, and more a seminatural interaction with the displayed objects, by using controllers. This advantage had becomes a very immersive experience for an actor's senses.

However, after the first enthusiastic pilots, commercial success originally VR did not succeed, because the communing environment was uncomfortable, unrealistic, very expensive and required computing power that was unavailable at that time – it was too early.

The situation has been changed in the last 10 years, and many of the initial problems of this technology have been resolved, with the fact of the information technology grows up. The current market of such devices, e.g.: Oculus Quest & Rift [4], HTC Vive, Valve Index, Google Cardboard [14], offers an open architecture for developers, and as results VR becomes available to a wide audience of consumers. The open architecture of the equipment supports easier production of VR applications, and it allows for quick popularization of this type of solutions.

2. VR in education

The knowledge transfer is a condition for the progress of civilization, therefore any innovations being used in acquiring education are the catalyst for the development of human beings in general. The educational domain development has a positive impact on a range and methods of providing information [8].

In the digitization era, Virtual Reality has become a medium that can be used in many ways, and in particular it is another logical step in the computers use for educational purposes – the spectrum of possible applications is wide [1, 16]. No other medium will achieve such a level of immersion and contact with the virtual world as VR is.

One of the biggest problems in education being based on the traditional information lecture is the student involvement for a long process of teaching. Lack of students' commitment leads to:

- problems with the knowledge acquisition,
- poor interests of the topic,
- lower efficiency of teaching process,

and in consequence it results in negative experiences and even the lack of any tangible learning outcomes.

Each didactic tool that allows for more extensive contact with the topic being taught makes scholars more involved into the process and they are more willing to analyse an issue, an generally it results onto positive effects of they own knowledge increase.
In this area the Virtual Reality enables a practical and extended contact with an issue, the internal interaction with virtual displayed objects in VR's applications is intuitive, and students are

willing to contact within virtual representation of a teaching challenge, due to its modern and attractive looks, and the relatively short period of its presence on the market. Therefore VR teaching applications strengthens students' involvement, allowing the next perspective of education to achieve the next level of sense interaction, just like in video games [2, 3].

2.1. Distance learning, visualization of complex environments

The traditional form of teaching, based on direct contact between students and the master, in many situations turns out to be insufficient or even impossible to be implemented. The bright example of such situation are the recent months, when due to the epidemiological situation around the world, almost all teaching activities had to be transferred into online platforms; but also as well before the pandemic there had been many cases where the availability of individual specialists' knowledge push the student's to leave their homes/cites and goes even out of the countries.

Virtual Reality allows us to move wherever we want, as well as interact with other users, if only the application was designed in this way. This adds variety of possibilities to the distance learning experience that can be realized even with a simple smartphone help for each students.

The weaknesses of the traditional didactic form also manifest themselves if the environment in which the examiners have be trained turns out to be dangerous, or it would even endanger the students' lives, and sometimes also if the training costs have no positive economic returns. The organization of educational courses, where the subject of which would be, for example: saving miners from a collapsed mine or utilization of life-threatening materials is practically impossible to implement in their real conditions – of course if such situation happens – the services guards have to show their own experience and in fact they increase their own experience to. But how about the first job, the first mission – when any work in such a difficult environment is practically impossible without exercise beforehand, or in the simplest case it is very expensive and risky for the trainee himself. For Virtual Reality these limitations do not exist, and the form of realization and customisation of a given environment depends only on the creators of a given virtual space. Currently, an extensive field in which VR is widely used is medicine – operations and procedures in which there is no agreement for the slightest error can be practiced and repeated many times using the Virtual Reality, and in this way prepare the future doctors for real operations [6].

2.2. Examples of mono-scenario VR teaching applications

In 2019, the Central Mining Institute has implemented the project titled "Specialized competences of the graduates as a chance for employment in the construction industry on the cross-border labour market". This project has assumed preparation and delivery of a virtual reality workshop, where the aim of it was to train users in the removal of hazardous materials by following the scenarios presented to them via a network connection. The application made it possible to gain practical experience in the removal of life-threatening materials (asbestos) without physical contact with them. The advantage of the application was the possibility of conducting training with two users at the same time through a network connection, which allows for training more people at the same time, and also prepares trainees for cooperation. The examiners were supervised by an administrator who could run scenarios, watch users' actions, as well as indicate errors or give commands on a separate computer. The workshop prepared in that project was mobile and could be assembled and set up anywhere. One of the authors of the current article has had the pleasure to participate in developing process of the above solution.

The limitation of that solution was the lack of the possibility of external interference with in the course of the scenario. The administrator could not change the characteristics of the items/tasks used by the trainees, which limits the functionality of the application. Anyway, this implementation is still actively used for user training [11, 12].

Another implementation that allows for only one scenario of interaction is the "Mission: ISS" application available from the Oculus Store [13]. It is an entertainment application that allows an user to play the role of a crew member of the International Space Station and explore its interior. However, the authors of the "Mission: ISS" went a step further than only an entertainment, and they faithfully have reproduced each element of the station, so they give the user a virtual document in which the former station members during a virtual tour tell about everyday life at the space station, describing its important elements, and also drawing the user's attention onto how their life looks like in the space. The application is free and available to anyone with the appropriate equipment. Thanks to this solution, the player is able to move to a place practically inaccessible to him, directly from his room, and thanks to the appropriate preparation of the application by the creators, he can gain knowledge from this experience that may seem boring. A virtual journey to such an exotic place, practically inaccessible to the common student, makes the experience extremely attractive, sharpens the senses and engages the participant to "absorb knowledge" somehow by the way. The disadvantage of the application is its linear structure and relatively small amount of content. After exploring once, the user may not be prompted to return to this application. The game may cause discomfort for some users, because no mechanics were used to increase the quality of the experience [5].

These two advantage applications shows, that even very realistic and attractive visualisation or even very important train course to be passed does not make an user to return back to the VR if it offers only one possible scenario of interactions. At least an user will be bored or will know the scenario as well that he/she will pass the exam without considering the current course. The authors assume that there is necessity to implement multi-scenario interaction with a VR teaching application, and such experiment will be described in the next chapters.

3. Multi-scenario VR application for logistics

The logistic warehouses are the space where many people starts their first job. At the first look young man doesn't need to high competitions to work physically there, but in this area there are many dangerousness as for the workers as well for the products being stored there. In this chapter there are described three typical situations which should be tested for warehouse workers' applicants, as in the examples: products confection; personal protective clothes use; safe path.

3.1. VR mechanic of products confection test

The first task for the player is to proper stack boxes onto appropriate pallets within a specified time and according to an order. Here any mistakes could results with wrong customers' orders confection and future reclamations or could results in the products damage if the boxes will not be fitted correctly.

Fig. 1. The products confection task [source: own]

The task itself does not have to consist of a fixed number of pallets, as an example, there are used 16 different boxes palettes and 4 pallets (see Fig. 1). The application is designed in a way that allows the developer to adapt the number of variables to a given variant of the task – it requires recompilation of the program, but it is enough to use the prepared development tools (described in subchapter 3.5) – it does not require interference with the source code of the application itself.

After the task start, the student "physically" grabs the boxes and moves them to the appropriate pallet, he use two controllers (see Fig. 7). When the box is on the appropriate pallet and it is stable (i.e. standing still), he can starts to load the next one. After the proper boxes arranging on a given pallet, the VR mechanic will show him the correct completion of the task by sublighting the pallet in green (see Fig. 2).

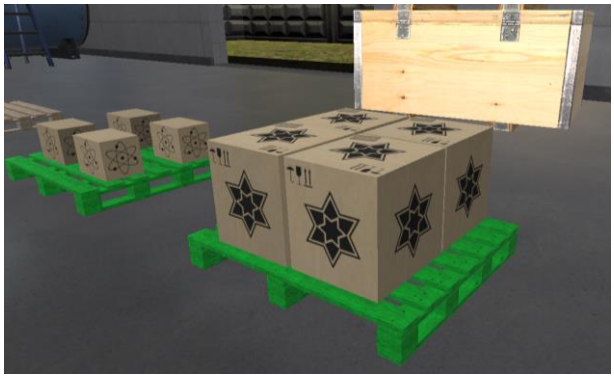


Fig. 2. The proper result of a task [source: own]

If, after placing some boxes on a pallet, it still shows red colour, it means that the student has done a mistake and must correct it. Some typical problems which could happen: objects protruding too much beyond the perimeter of the pallet, inappropriate items are put on the pallet or careless positioning of the item itself. In each case, the student have to correct his mistake by approaching the pallet and moving the boxes to the appropriate places.

3.2. VR mechanic of protective gloves test

The second type of mechanics which could be added to the scenario is the task in which a student has to choose the appropriate gloves for cleaning the workplace. There are sharp objects there – broken glass, so before starting work, he has to choose the right protective gloves for this task.

Fig. 3. The appropriate gloves used to a cleaning test [source: own]

Out of the 3 types of gloves, two are correct, but only one type is provided in the workplace (e.g. the leather gloves are correct and available, rubber ones are available too but not correct, the firefighting gloves are safe and available, but they are a specialized equipment for special use and the worker shouldn't use it for casual work).

In this mechanics, the user's prior preparation for training is tested and he will pass the test after the correct choice. Gloves are represented by the appropriate material being put on the student's hands in contact with the prepared objects (see Fig. 3).

3.3. VR mechanic of safe path test

The third type of mechanics which could be added to the scenario is a passive mechanics, where a student has to remember about his own safety. This mechanic continues from the test beginning to its end, and the user has to keep careful and remember about necessity of the safety path, because at the same warehouse there are many forklifts and the workers cannot walk everywhere – it is too dangerous. If the mechanic is added to the test, the task consists of paths / communication routes that are used by forklifts to move the loads. The safety paths are marked by thin white stripes in the places where the examiner's entry is safe and withe the wide ones in the places where his move means an error (see Fig. 4).

Fig. 4. The wrong cross of a forklift communication path [source: own]

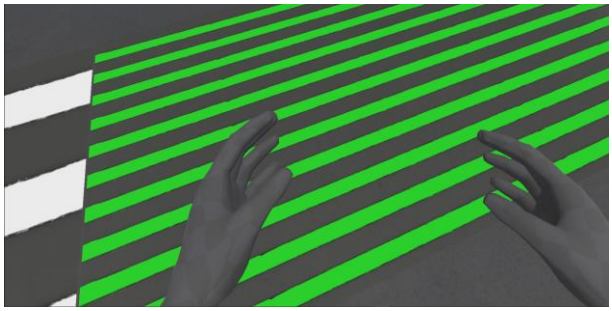


Fig. 5. The proper cross of a forklift communication path [source: own]

Those crosses of communication routes are usually marked and the operators of forklifts are required to increase vigilance in these places – so it is more safe to walk there, than in another places, were the drivers go more fast. If a student enters the wrong area – the place he has violated will immediately turn red to inform him about the error made. Similarly, when he enter the correct area – this will be highlighted in green, to inform him about the correct place of passage (see Fig. 5). Such visual anchors allow the user to get used to similar places in reality of a physical warehouse – a student after the training will remember that the densely placed strips are for his safety, and other belts – for vehicles located in the warehouse.

3.4. Multi-selection panel

The user is the entity that has the closest contact with the application – he puts the equipment on his head and manipulates virtual objects by performing specific activities.

The above tasks have been designed to be put into scenarios as full competition testing or as well in an individual selected parts for a dedicated user. The all prepared mechanics could be switched on or off on the main panel of the application (see Fig 6). The user is the entity that has the closest contact with the application – he puts the VR equipment on his head and manipulates virtual objects by performing specific activities. The mechanics are prepared to simulate some real situations in which the employee is supposed to demonstrate perceptiveness, orientation in the field in order to maintain safety, or select the appropriate tool for a given task.

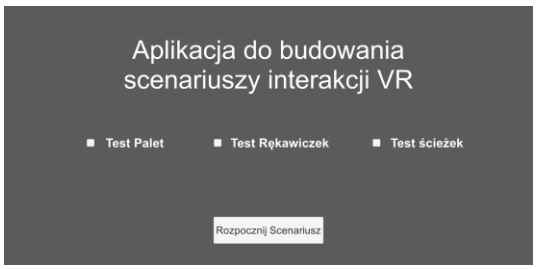


Fig. 6. The main panel layout [source: own]

That an administrator/examiner decides which ones should be passed by the tested user.

3.5. Software and technological background

The equipment for which the application has been designed is a modern and up-to-date Virtual Reality Oculus Quest [14] goggles. They are equipped with two controllers, a head set and a cable connecting the equipment to the computer (see fig. 7).

The native operating environment of this model of goggles are the internal peripherals of the device itself (built-in components that allow you to run the application without the need for any external hardware), but thanks to the Oculus Link technology [15] In the case of cable connection with a computer, the goggles can use its computing power to more advanced applications. The prepared solution assumes the use of this option.

Fig. 7. Oculus layout [source: www.oculus.com]

The solution was made using the modern UnityXR framework, allowing the program to be compiled on any VR device without the need to adapt it to hardware differences between platforms. It is a relatively new solution that significantly unifies the way of creating similar applications between different platforms (the way actions are performed on controllers, traffic recognition, etc.). As a result, the process of creating applications for various VR platforms has been significantly simplified, because it is not necessary to spend a lot of time adjusting the solution to the equipment of different manufacturers.

The engine that was used to build and compile the program is Unity, a complete and supported for years solution for creating games, commercial and educational programs, as well as animations. The engine version that was used to build the solution is 2019.4.1f1, but thanks to the use of UnityXR, which is a built-in tool in the engine, it is possible to update the compiler version.

The code editor that was used to build the solution is Visual Studio Community 2019, a common and proven solution that can be natively connected to Unity, allowing for efficient work and code debugging. The mechanics and the entire software logic have been written in C#, which is a standard solution when working with the Unity engine.

4. Conclusions

The 3D scenes prepared in the presented VR training application are not such effective as e.g. presented in the Mission: ISS application, but that wasn't the main goal of the authors. Using simple graphics we attempt to prove the hypothesis that the multi-variant interaction of a user with the virtual world is useful. In the article we describe the sample VR training application for a logistics warehouse worker.

To create a dedicated training scenario in the Unity environment, it does not relay only on the layer of manipulation of windows in this environment. In order to be able to navigate through the previously developed multi-variant scenarios, it is necessary to prepare a script that activates individual scenarios. Such scripts being prepared properly in C # can be added to any element visible on the stage, regardless of whether it is a graphic object, UI part or even an empty GameObject, and give them the characteristics that have been defined through the code.

Thus, it has been shown that the user's interaction with the virtual world can be personalized to the needs of a given training scenario, ant it would be more useful than repeats the same activities over and over again. The assumptions of the authors' work were achieved and the hypothesis has been proved.

References

[1] Çankaya S.: Use of VR Headsets in Education: A Systematic Review Study. Journal of Educational Technology and Online Learning 2(1)/2019, 74–88.
[2] Delialioglu O.: Student Engagement in Blended Learning Environments with Lecture-Based and Problem-Based Instructional Approaches. Educational Technology & Society 15(3)/2012, 310–322.
[3] Hu-Au E., Lee J.: Virtual reality in education: a tool for learning in the experience age. International Journal of Innovation in Education 4(4), 2017.
[4] Kent S.: The Ultimate History of Video Games: The Story Behind the Craze that Touched our Lives and Changed the World. Random House, 2001.
[5] Norouzi N., Bruder G., Welch G.: Assessing vignetting as a means to reduce VR sickness during amplified head rotations. Proceedings of the 15th ACM Symposium on Applied Perception (SAP'18) 19/2018, 1–8.
[6] Riener R., Harders M.: Virtual Reality in Medicine. Springer, London 2012.
[7] Satava R. M.: Virtual reality, telesurgery, and the new world order of medicine. Journal of Image Guided Surgery 1(1)/1995, 12–16.
[8] Skibska J., Wojciechowska J.: Współczesna edukacja. Wielopłaszczyznowość zadań. Wydawnictwo LIBRON, Kraków 2016.
[9] Steinicke F.: Being Really Virtual. Immersive Natives and the Future of Virtual Reality. Springer, 2016.
[10] Vorländera M., Schrödera D., Pelzera S., Wefersaa F.: Virtual reality for architectural acoustics. Journal of Building Performance Simulation 8(1), 15–25.
[11] https://www.gig.eu/pl/newsy/spotkanie-otwierajace-projekt-pt-specjalistyczne-kompetencje-absolwentow-szansa-na (available: 10.01.2020).
[12] https://www.gig.eu/pl/przetargi/aktualne/przygotowanie-i-dostarczenie-pracowni-wirtualnej-rzeczywistosci (available: 10.01.2020).
[13] https://www.oculus.com/quest/features/ (available: 21.01.2021).
[14] https://arvr.google.com/cardboard/ (available: 11.01.2021).
[15] https://www.pocket-lint.com/ar-vr/news/oculus-rift/153723-what-is-oculus-link-and-how-do-you-use-it (available: 11.09.2020).
[16] https://xd.adobe.com/ideas/principles/emerging-technology/virtual-reality-will-change-learn-teach/ (available: 19.08.2020).

Eng. Wojciech Włodyka
e-mail: wlodykawojciech@gmail.com

Engineer at the University of Rzeszów, Enthusiast of Virtual Reality and information technologies.

http://orcid.org/0000-0002-8015-9657

Ph.D. Eng. Dariusz Bober
e-mail: dbober@ur.edu.pl

Assistant professor at University of Rzeszów, Department of Computer Science, specialist in databases, industrial system, logistics and electrical power management. The creator of three patents and the Power Modes methodology of energy management

http://orcid.org/0000-0002-0480-700X

otrzymano/received: 21.02.2021 *przyjęto do druku/accepted: 15.03.2021*

http://doi.org/10.35784/iapgos.2404

INVESTIGATION OF THE DEPENDENCE OF THE STRUCTURE OF SHIFT INDEXES VECTORS ON THE PROPERTIES OF RING CODES IN THE MOBILE NETWORKS OF THE INTERNET OF THINGS

Vladislav Kravchenko, Olena Hryshchenko, Viktoriia Skrypnik, Hanna Dudarieva
State University of Telecommunications, Kiev, Ukraine

Abstract. *The essence of the concept of a family of several codes and its properties is considered. Typical structures of shear vectors for a family of several codes are analyzed. The regularity of determining the decimal values of the elements of the shift vectors and their dependence on the length of the code combinations and the number of single characters in the code combination is determined. The defined regularities allow algorithms to be developed to decode information with the use of the structure of the shift vectors.*

Keywords: ring code, ring code family, shift indexes vector, code combination length

BADANIE ZALEŻNOŚCI STRUKTURY WEKTORÓW INDEKSÓW PRZESUNIĘCIA OD WŁAŚCIWOŚCI KODÓW PIERŚCIENIOWYCH W MOBILNYCH SIECIACH INTERNETU RZECZY

Streszczenie. *Rozważono istotę pojęcia rodziny kodów pierścieniowych i jej właściwości. Przeanalizowano typowe struktury wektorów przesunięć dla rodziny kilku kodów. Określono prawidłowości wyznaczania wartości dziesiętnych elementów wektorów przesunięć i ich zależności od długości kombinacji kodowych oraz liczby pojedynczych znaków w kombinacji kodowej. Wyznaczone prawidłowości pozwalają na opracowanie algorytmów dekodowania informacji z wykorzystaniem struktury wektorów przesunięcia.*

Słowa kluczowe: kod pierścieniowy, rodzina kodów pierścieniowych, wektor offsetowy, długość kombinacji kodów

Introduction

According to [1–3], ring codes are built on the principle block cyclic code rows forming matrices that are interconnected by the condition of cyclicity. The ring code differs from the cyclic code in that the generating matrix of the ring code is always square and contains N columns and N rows, each row of which consists of m ones and N - m zeros [1–3].

Ring codes are characterized by a vector of shear shift which is formed by adding the number of units obtained by bitwise implementation of one of the binary transformations XOR, OR, AND elements of the first line of the forming matrix of the ring code and the rest of its lines [3]. The shear vector shift is a group integral index of the ring code and the number of elements in the vector of the shear shift is always one less than the number of elements in the code combination of the ring code.

Any vector of shear shift characterizes a family of ring codes of a certain type.

1. Properties of ring code families

Each line of the generating matrix of the ring code is characterized by a delta factor – the distribution of zero and single symbols between the two extreme units, separated by the largest number for this initial vector of zero symbols. Ring codes that have a delta factor of a certain type create a family of ring codes. For example, Table 1 shows a family of ring codes depending on the type of delta factor

Table 1. Families of ring codes depending on the type of delta factor

Structure of code combinations of the family of ring codes of type 000111	Structure of code combinations of the family of ring codes of type 010101	Structure of the shear vector
0000000001	0101	0000000001
0000000011	010101	0000000101
0000000111	01010101	0000001101
0000001111	0101010101	0000001101
0000011111	010101010101	0000011101
0000111111	01010101010101	0000111101
0001111111	0101010101010101	0001111101

The first column of Table 1 shows a family of ring codes of type 000111, single symbols in the delta factor that are placed in a row. The second column of the Table shows a family of ring codes of type 010101, single and zero symbols in the delta factor that alternate with each other. The third column of the Table shows a family of ring codes of type 001101, in the delta factor of which the ones are placed both in a row and alternating with zero symbols. Thus, within a family of codes, the ring structure combinations of the ones and zeros in the code combinations are identical. Code combinations differ only in the number of ones and zeros, without changing the structure of the combinations of the ones and zeros.

2. Structure of vector of a shear shift of ring codes

In order to analyze the structure of the vectors of the shift of the ring codes, the vectors of the shear shift for three types of families of ring codes are considered: type 000111, type 010101 and type 010011. The study of the structure of the shear vectors of the above families allows us to note that within a certain family, the vectors of the shear shift have the same structure of alternation of decimal values of elements. The value of the decimal value of the element of the shear shift vectors depends on both the length of the code combination N and the number of unit symbols m and zero symbols N - m in the code combination of the ring code. The number of decimal values in the vector of the shear indicators is always one less than the number of elements in the code combination, and the number of code combinations of the ring code.

2.1. Analysis of the structure of vectors of a shear shift of ring codes of family type 000111

Table 2 shows the dynamics of the structure of the vectors of the shear shift depending on the length of the code combination *N* with a constant number of units m in the code combination of ring codes of the family 000111. Thus in 2 columns of Table 2 the first line of the forming matrix of a ring code is given.

Table 2. Dynamics of changes in the structure of the vector of the shear shift of the ring code type 000111 depending on the length of the code combination N at m = 3

Length of the code combination N	Structure of the code combination	Structure of the shear vector shift
7	0000111	246642
8	00000111	2466642
9	000000111	24666642
10	0000000111	246666642
11	00000000111	2466666642
12	000000000111	2466666642

Table 3 shows the dynamics of changes in the structure of the vectors of the shear shift depending on the number of units m in the code combination with a constant length of the code combination N ring codes.

Table 3. Dynamics of changes in the structure of the vector of the shear shift of the ring code type 000111 depending on the number of units m at N = 8

Number of units of m	Structure of the code combination	Structure of the shear vector shift
1	00000001	2222222
2	00000011	2444442
3	00000111	2466642
4	00001111	2468642
5	00011111	2466642
6	00111111	2444442
7	01111111	2222222

Figure 1 shows a graph of the values of the decimal values of the elements D of the shear vector shift from the position number of the element in the vector of the shear shift P for a family of ring code type 000111 length N = 8 for a different number of unit symbols m.

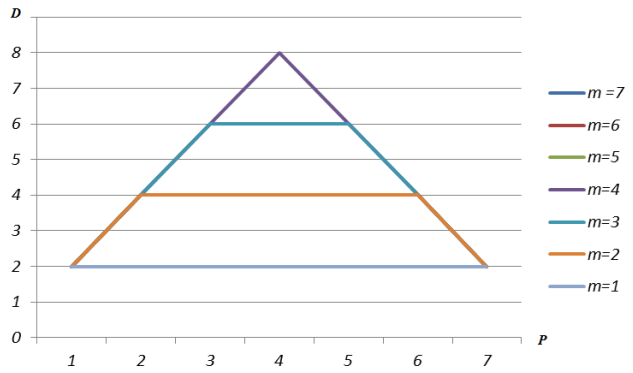


Fig. 1. Graph of the dependence of the value of the decimal values of the elements of the vector of the shear shift on the position number of the element for the code combination of the family 000111

An analysis of the dynamics of the change of the structure of the vectors of the shear shift of the ring code of type 000111, depending on the length of code combination N and the quantity of units m given in Tables 2,3 and the dependence of the values of the decimal values of elements D of the vectors of the shear shift on the number of the position of the placement of the element in the vector of shift P, presented in Fig. 1, allows us to note that the numerical series of decimal values of the vector of the shift begins and ends with the number 2.

Also, the numerical series of decimal values of the vector of shear shift consists of several sets, depending on the number of units in the code sequence and the ratio of the number of units and the number of zeros.

Thus, under the condition m = 1 or N - m = 1, a numerical series of length N - 1 = 7 elements is formed, the numerical values of which are equal to 2m, which means 2.

If m > 1 or N - m > 1, a numerical series is formed, which consists of three sets:

$$\bar{B} \rightarrow S_1 \cup S_2 \cup S_3 \quad (1)$$

The decimal values of the elements in the set S_1 vary in arithmetic progression from 2 to 2m and the number of such elements k1 is equal to:

$$k_1(m)=\begin{cases} 2m, m \leq N-m; \\ 2(N-m), N-m \leq m. \end{cases} \quad (2)$$

The decimal values of the elements in the set S_2 do not change and are equal to $2m$ and the number of such elements k_2 is equal to:

$$k_2(m)=\begin{cases} N-m, m < N-m; \\ N-(N-m), m > N-m. \end{cases} \quad (3)$$

The decimal values of the elements in the set S_3 change into arithmetic progressions from 2m - 1 to 2.

2.2. Analysis of the structure of vectors of a shear shift of ring codes of family type 010101

Table 4 shows the dynamics of the structure of the vectors of the shear shift depending on the length of the code combination N of the ring codes of the family 010101. Thus, in 2 columns of the table, the first line of the forming matrix of a ring code is given.

Table 4. Dynamics of change of structure of vectors of the shear shift

Length of code combination N	Structure of code combination	Structure of shear vector shift
4	0101	404
6	010101	60606
8	01010101	8080808
10	0101010101	10 0 10 0 10 0 10 0 10
12	010101010101	12 0 12 0 12 0 12 0 12 0 12
14	01010101010101	14 0 14 0 14 0 14 0 14 0 14 0 14

Table 5 shows the dynamics of the changes in the structure of the vectors of the shear shift depending on the number of units m in the code combination of the ring codes of the family 010101.

Table 5. Dynamics of changes in the structure of the vectors of shear shift ring code type 010101 depending on the number of units m

Number of units m	Structure of code combination	Structure of shear vector shift
2	0101	404
3	010101	60606
4	01010101	8080808
5	0101010101	10 0 10 0 10 0 10 0 10
6	010101010101	12 0 12 0 12 0 12 0 12 0 12
7	01010101010101	14 0 14 0 14 0 14 0 14 0 14 0 14

Figure 1 shows a graph of the values of the decimal values of elements D of the shear vector shift from the position number of the element in the vector of shear shift P for the family of ring code type 010101 for a different number of unit symbols m.

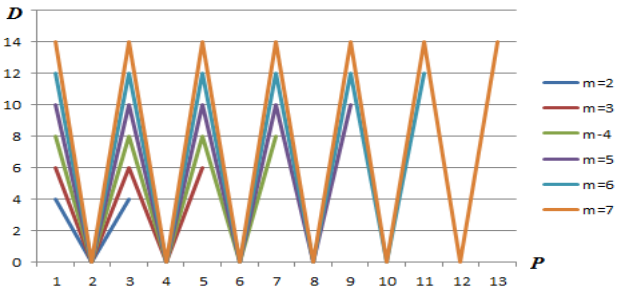


Fig. 2. Graph of the dependence of the value of the decimal values of the elements of the vector of the shear shift on the position number of the element for the code combination of the family 010101

The analysis of the dynamics of changes in the structure of the shear vector shift of the ring code type 010101, depending on the length of the code combination N and the number of units m shown in Table 4 and 5 and the dependence of the decimal values of the elements D of the shear vector shift on the position number of the element in the shear vector shift P, presented in Figure 2, allows us to note that the numerical series of decimal values of the vector of the shear shift begins and ends with a number equal to N = 2m.

In this case, as can be seen from Table 4 and 5, the family of ring code type 010101 is characterized by code combinations with an even number of elements and when changing both the length N of the code combination and the number of units m in the code combination, the structure and dynamics of changing the structure of the vectors of the shear shift are the same. Therefore, there is an unambiguous dependence of the structure of the vectors of the shear shift on both the length N of the code combination and the number of units m in each code combination. The structure of the vectors of the shear shift is not repeated when changing the number of units in the code combination, which allows us to uniquely determine the code combination on the vector of the shear shift.

The number of elements k3 of the vector of the share shift of the ring code family type 010101 is equal to:

$$k_3(m)=\begin{cases} m, x_k = 2m; \\ m-1, x_k = 0, \end{cases} \quad (4)$$

where: x_k – decimal value of the element vector of the shear shift.

2.3. Analysis of the structure of vectors of a shear shift of ring codes of family type 001101

Table 6 shows the dynamics of the structure of the vectors of the shear shift depending on the length of the code combination with a constant number of units m in the code combination of the ring codes of the family 001101. Thus in 2 columns of the table, the first line of the forming matrix of a ring code is given.

Table 6. Dynamics of changes in the structure of the vector of the shear shift of the ring code type 001101 depending on the length of the code combination N at m = 3

Length of code combination N	Structure of code combination	Structure of vector of shear shift
5	01101	4224
6	001101	44244
7	0001101	444444
8	00001101	4446444
9	000001101	44466444
10	0000001101	444666444
11	00000001101	4446666444
12	000000001101	44466666444

Table 7 shows the dynamics of changes in the structure of the vectors of the shear shift depending on the number of units m in the code combination with a constant length of the code combination of N ring codes.

Table 7. Dynamics of changes in the structure of the vectors of the shear shift of ring code type 001101 depending on the number of units m at N = 9

Number of units m	Structure of code combinations	Structure of the vector of shear shift
2	000000101	42444424
3	000001101	44466444
4	000011101	44666644
5	000111101	44666644
6	001111101	44466444
7	011111101	42444424

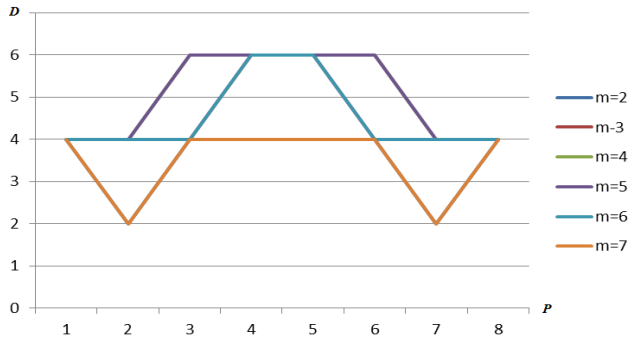


Fig. 3. Graph of the value of the decimal values of the elements of the vector of the shear shift from the position number of the element for the code combination of the family 001101

An analysis of the dynamics of changes in the structure of the VPZ ring code type 001101, depending on the length of the code combination N and the number of units m shown in Tables 6 and 7 and the dependence of the values of the decimal values of the elements of D vector of the shear shift from the position number of the elements presented in Fig. 3, shows that the numerical series of decimal values of the vector of the shear shift begins and ends with the number 4. The maximum decimal value of the element of the vector of the shear shift is equal to 2m, meaning equal to 6. However, the unambiguous dependence of the structure of the vectors of the shear shift on the number of unit symbols m is not observed. In addition, the structure of the vectors of the shear shift is repeated when changing the number of units in the code combination, which does not allow the code combination to be unambiguously determined by the vector of the shear shift.

3. Conclusions

The analysis of the dependence of the structure of the vector shear shift on the properties of ring codes allows us to note that different families of ring codes correspond to different structures of the vector shear shift. Within each family of ring codes, the identity of the values of the decimal values of the elements of the vectors of the shear shift, their combinations and dynamics of their change, is observed. The dependence of the values of the decimal values of the elements on the length N of the code combinations, the number of unit elements m and the ratio of ones and zeros in the code combination is observed.

The most pronounced abovementioned dependence is observed for the family of type 010101, and the least pronounced dependence is observed for the family of type 001101.

References

[1] Berkman L., Otrokh S., Kuzminykh V., Hryshchenko O.: Method of formation shift indexes vector by minimization of polynomials. CEUR Workshop Proceedings 2577, 2019, 259–269.
[2] Gavrilko E. V., Otrokh S. I., Yarosh V. A., Grishchenko L. N.: Improving the quality of the functioning of the network of the future through the use of ring codes. Vesnik svyazi 2/2018, 60–64.
[3] Otrokh S., Kuzminykh V., Hryshchenko O.: Method of forming the ring codes. CEUR Workshop Proceedings 2318, 2018, 188–198.

Ph.D. Vladislav Kravchenko
e-mail: vladkr58@gmail.com

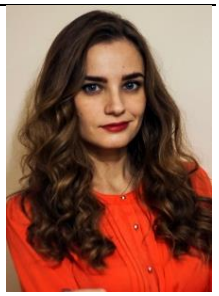
Head of the Department of Mobile and Video Information Technologies of the State University of Telecommunications, Kiev, Ukraine. Research interests: LTE, OFDM, information security.
Publications: about 35 scientific publications, including 3 patents.

http://orcid.org/0000-0002-4758-7027

M.Sc. Olena Hryshchenko
e-mail: elena.grischenko1@gmail.com

Graduate student of the State University of Telecommunications, Kiev, Ukraine. Research interests: information security, information coding.
Publications: about 8 scientific publications.

http://orcid.org/0000-0001-8198-5056



M.Sc. Viktoriia Skrypnik
e-mail: skripnik2008@ukr.net

Graduate student of the State University of Telecommunications, Kiev, Ukraine. Research interests: information security, information coding.
Publications: about 6 scientific publications.

http://orcid.org/0000-0001-9854-8039

M.Sc. Hanna Dudarieva
e-mail: annett.13.86@gmail.com

Graduate student of the State University of Telecommunications, Kiev, Ukraine. Research interests: information security, information coding.
Publications: about 7 scientific publications.

http://orcid.org/0000-0002-9887-021X

otrzymano/received: 23.11.2020 *przyjęto do druku/accepted: 15.03.2021*