p-ISSN 1607-3274
e-ISSN 2313-688X

# Радіоелектроніка Інформатика Управління

# Radio Electronics Computer Science Control

2024/4

Національний університет «Запорізька політехніка»

# Радіоелектроніка, інформатика, управління

Науковий журнал
Виходить чотири рази на рік
№ 4(71) 2024
Заснований у 1998 році, видається з 1999 року.
Засновник і видавець – Національний університет «Запорізька політехніка».
ISSN 1607-3274 (друкований), ISSN 2313-688X (електронний).
Запоріжжя
НУ «Запорізька політехніка»
2024

National University Zaporizhzhia Polytechnic

# Radio Electronics, Computer Science, Control

The scientific journal
Published four times per year
№ 4(71) 2024
Founded in 1998, published since 1999.
Founder and publisher – National University Zaporizhzhia Polytechnic.
ISSN 1607-3274 (print), ISSN 2313-688X (on-line).
Zaporizhzhia
NU Zaporizhzhia Polytechnic
2024

**Науковий журнал «Радіоелектроніка, інформатика, управління»** (скорочена назва – РІУ) видається Національним університетом «Запорізька політехніка» (НУ «Запорізька політехніка») з 1999 р. періодичністю чотири номери на рік.

Реєстрація суб'єкта у сфері друкованих медіа: Рішення Національної ради України з питань телебачення і радіомовлення № 3040 від 07.11.2024 року. Ідентифікатор медіа: R30-05582.

**ISSN** 1607-3274 (друкований), **ISSN** 2313-688X (електронний).

Наказом Міністерства освіти і науки України № 409 від 17.03.2020 р. «Про затвердження рішень Атестаційної колегії Міністерства щодо діяльності спеціалізованих вчених рад від 06 березня 2020 року» **журнал включений до переліку наукових фахових видань України в категої «А» (найвищій рівень),** в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора наук і доктора філософії (кандидата наук).

Журнал включений до польського Переліку наукових журналів та рецензованих матеріалів міжнародних конференцій з присвоєною кількістю балів (додаток до оголошення Міністра науки та вищої освіти Республіки Польща від 31 липня 2019 р.: № 16981).

В журналі безкоштовно публікуються наукові статті англійською, російською та українською мовами.

**Правила оформлення статей** подано на сайті: http://ric.zntu.edu.ua/information/authors.

Журнал забезпечує **безкоштовний відкритий он-лайн доступ** до повнотекстових публікацій.

Журнал дозволяє авторам мати авторські права і зберігати права на видання без обмежень. Журнал дозволяє користувачам читати, завантажувати, копіювати, поширювати, друкувати, шукати або посилатися на повні тексти своїх статей. Журнал дозволяє повторне використання його вмісту у відповідності Creative Commons ліцензією CC BY-SA..

Опублікованим статтям присвоюється унікальний ідентифікатор цифрового об'єкта DOI.

**Журнал входить до наукометричної бази Web of Science.**

**Журнал реферується та індексується** у провідних міжнародних та національних реферативних журналах і наукометричних базах даних, а також розміщується у цифрових архівах та бібліотеках з безкоштовним доступом у режимі on-line, повний перелік яких подано на сайті: http://ric.zntu.edu.ua/about/editorialPolicies#custom-0.

**Журнал розповсюджується** за Каталогом періодичних видань України (передплатний індекс – 22914).

**Тематика журналу**: телекомунікації та радіоелектроніка, програмна інженерія (включаючи теорію алгоритмів і програмування), комп'ютерні науки (математичне і комп'ютерне моделювання, оптимізація і дослідження операцій, управління в технічних системах, міжмашинна і людино-машинна взаємодія, штучний інтелект, включаючи системи, засновані на знаннях, і експертні системи, інтелектуальний аналіз даних, розпізнавання образів, штучні нейронні і нейро-нечіткі мережі, нечітку логіку, колективний інтелект і мультиагентні системи, гібридні системи), комп'ютерна інженерія (апаратне забезпечення обчислювальної техніки, комп'ютерні мережі), інформаційні системи та технології (структури та бази даних, системи, засновані на знаннях та експертні системи, обробка даних і сигналів).

Усі статті, пропоновані до публікації, одержують **об'єктивний розгляд**, що оцінюється за суттю без урахування раси, статі, віросповідання, етнічного походження, громадянства або політичної філософії автора(ів).

Усі статті проходять двоступінчасте закрите (анонімне для автора) **резензування** штатними редакторами і незалежними рецензентами – провідними вченими за профілем журналу.



Рекомендовано до видання Вченою радою НУ «Запорізька політехніка», протокол № 4 від 26.11.2024.

Журнал зверстаний редакційно-видавничим відділом НУ «Запорізька політехніка».

**Веб-сайт журналу**: http://ric.zntu.edu.ua.

**Адреса редакції:** Редакція журналу «РІУ», Національний університет «Запорізька політехніка», вул. Жуковського, 64, м. Запоріжжя, 69063, Україна.

Тел: (061) 769-82-96 – редакційно-видавничий відділ Факс: +38-061-764-46-62

E-mail: rvv@zntu.edu.ua

**The scientific journal Radio Electronics, Computer Science, Control** is published by the National University Zaporizhzhia Polytechnic NU Zaporizhzhia Polytechnic since 1999 with periodicity four numbers per year.

Registration of an entity in the field of print media: Decision of the National Council of Ukraine on Television and Radio Broadcasting No. 3040 of November 7, 2024. Media ID: R30-05582.

**ISSN** 1607-3274 (print), **ISSN** 2313-688X (on-line).

By the Order of the Ministry of Education and Science of Ukraine from 17.03.2020 № 409 "On approval of the decision of the Certifying Collegium of the Ministry on the activities of the specialized scientific councils dated 06 March 2020" **journal is included in the list of scientific specialized periodicals of Ukraine in category "A" (highest level),** where the results of dissertations for Doctor of Science and Doctor of Philosophy may be published.

The journal is included to tne Polish List of scientific journals and peer-reviewed materials from international conferences with assigned number of points (Annex to the announcement of the Minister of Science and Higher Education of Poland from July 31, 2019: Lp. 16981).

The journal publishes scientific articles in English, Russian, and Ukrainian free of charge.

The **article formatting rules** are presented on the site: http://ric.zntu.edu.ua/information/authors.

The journal provides policy of **on-line open (free of charge) access** for full-text publications. The journal allow the authors to hold the copyright without restrictions and to retain publishing rights without restrictions. The journal allow readers to read, download, copy, distribute, print, search, or link to the full texts of its articles. The journal allow reuse and remixing of its content, in accordance with Creative Commons license CC BY-SA.

Published articles have a unique digital object identifier (DOI).

**The journal is included into Web of Science.**

**The journal is abstracted and indexed** in leading international and national abstractig journals and scientometric databases, and also placed to the digital archives and libraries with a free on-line access, full list of which is presented at the site: http://ric.zntu.edu.ua/about/editorialPolicies#custom-0.

**The journal is distributed**:by the Catalogue of Ukrainian periodicals (the catalog number is 22914).

**The journal scope:** telecommunications and radio electronics, software engineering (including algorithm and programming theory), computer science (mathematical modeling and computer simulation, optimization and operations research, control in technical systems, machine-machine and man-machine interfacing, artificial intelligence, including data mining, pattern recognition, artificial neural and neuro-fuzzy networks, fuzzy logic, swarm intelligence and multiagent systems, hybrid systems), computer engineering (computer hardware, computer networks), information systems and technologies (data structures and bases, knowledge-based and expert systems, data and signal processing methods).

All articles proposed for publication receive an **objective review** that evaluates substantially without regard to race, sex, religion, ethnic origin, nationality, or political philosophy of the author(s).

All articles undergo a two-stage **blind peer review** by the editorial staff and independent reviewers – the leading scientists on the profile of the journal.

## EDITORIAL BOARD



## EDITORIAL-ADVISORY COUNCIL



Recommended for publication by the Academic Council of NU Zaporizhzhia Polytechnic, protocol № 4 dated 26.11.2024.

The journal is imposed by the editorial-publishing department of NU Zaporizhzhia Polytechnic.

**The journal web-site** is http://ric.zntu.edu.ua.

**The address of the editorial office:** Editorial office of the journal Radio Electronics, Computer Science, Control, National University Zaporizhzhia Polytechnic, Zhukovskiy street, 64, Zaporizhzhia, 69063, Ukraine.

Tel.: +38-061-769-82-96 – the editorial-publishing department. Fax: +38-061-764-46-62

E-mail: rvv@zntu.edu.ua

# ЗМІСТ

# CONTENTS

p-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2024. № 4
e-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2024. № 4

# РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

# RADIO ELECTRONICS AND TELECOMMUNICATIONS

UDC 621.396.946

## IDENTIFICATION OF MOBILE DEVICES BY CORRELATION FEATURES OF THEIR SIGNAL SPECTRA

**Antipov I. –** Dr. Sc., Professor, Associate Professor of the Department of Computer Radio Engineering and Technical Information Protection Systems, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

**Vasylenko T.** – PhD, Senior Lecturer of the Department of Computer Radio Engineering and Technical Information Protection Systems, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

### ABSTRACT

**Context.** The mass spread of Wi-Fi networks is facilitated by the simplicity of their deployment, high speed, universality, and convenience of use. The development and dissemination of these networks continue despite a number of shortcomings. One of the shortcomings is their vulnerability to various types of attacks, including those based on the forgery (imitation) of identification data. At the same time, there are physical layer characteristics, knowledge of which expands the understanding of the network's state, can contribute to increasing the reliability of network subscriber identification, and thus prevent a number of attacks. This research is aimed at the theoretical and practical substantiation of the possibility of their application.

**Objective.** The aim of the study is to assess the application of detailed analysis of signal spectra emitted by devices connected to wireless Wi-Fi networks for their identification. To achieve this goal, it is necessary to analyze the experimentally measured spectra of wireless devices connected to the Wi-Fi network and evaluate the possibility of using the spectrum for the identification of mobile devices.

**Method.** This work proposes a method for processing the results of measuring the spectra of Wi-Fi device emissions by evaluating the asymmetry coefficient of the Wi-Fi device spectrum's cross-correlation function. Mathematical modeling was used to assess the effectiveness of the method.

**Results.** The research results show that the minimum value of the asymmetry coefficient when comparing the template with different positions of one's own device, and large values of the asymmetry coefficient when comparing templates with foreign spectra. Therefore, this characteristic can also be used for the identification of Wi-Fi devices.

**Conclusions.** The research results suggest the possibility of applying the proposed method for the identification of mobile devices, which will qualitatively complement existing security models with another feature for detecting unauthorized access.

**KEYWORDS:** security, Wi-Fi, identification, spectrum, asymmetry coefficient, mobile device.

### ABBREVIATIONS

A is a smartphone Redmi note 4X;

ACF is an autocorrelation function;

B is a smartphone Redmi note 4X, similar to "A", but the second copy;

C is a smartphone MeizuM5 Note;

CCF is a cross-correlation function;

D is a smartphone Honor 09 Lite;

E is a smartphone MeizuM6 Note;

IDS is an intrusion Detection System;

IoT is an Internet of Things;

MSD is a mean squared deviation;

OSI is a open system interconnection;

Wi-Fi is a wireless fidelity;

WPA is a Wi-Fi Protected Access.

### NOMENCLATURE

$A$ is a coefficient of asymmetry;

$B(\tau)$ is a degree of signal difference;

$B(j)$ is a correlation function;

$B_{cp}$ is an average value of the correlation function;

$i$ is a spectral component number;

$j$ is a difference in spectral components;

$K$ is a number of mobile users;

$m_3$ is a central empirical moment of the third order;

$N$ is a number of spectral components;

$P_{L1}(f_i)$ is a power of each spectral component;

$P_{L2}(f_i)$ is a power of each spectral component of the connecting subscriber;

$P_{L2}(f_{i+j})$ is a power of each spectral component displaced spectral components;

$S(t)$ is a harmonic signal;

$S(t–\tau)$ is a harmonic signal is shifted in time;

$\sigma$ is a mean squared deviation;

$\Delta$ is a spectrum width at the level of 0.5.

### INTRODUCTION

The rapid development of Wi-Fi networks covers all spheres of human activity. The principle of building wireless networks carries not only advantages in the form of free movement in the coverage area, sufficient data


DOI 10.15588/1607-3274-2024-4-1

transmission speed and low deployment cost, but also a lot of vulnerabilities and threats.

The main advantage of a Wi-Fi network is the transmission of information over radio waves without using wires, but at the same time it is the greatest threat to transmitted information, because controlling information transmitted through the air is not an easy task. Analysis of threats and attacks on wireless networks [1, 2] shows that most often for the removal or modification of information transmitted or stored in the network is the use of foreign equipment, which is very often disguised as network subscribers. In almost every type of attack there is an element that is disguised as a network subscriber.

To protect wireless networks from attacks, IDS is used. They are able to detect and prevent attacks by restricting access to the network or changing the configuration of communication equipment. Signs of attacks in existing IDS are the parameters of network traffic (node network activity, node network settings, data on files and processes) that is, the signs of channel, network and higher levels of the OSI model. This approach is fully justified in leading or fiber-optic networks, where physical connection to the network for intruders is difficult, and therefore the identification of equipment as such is absent (only user authentication is carried out). However, connecting to a Wi-Fi network at the physical level is not a problem for intruders through a torn radio interface.

Therefore, standard protection measures do not provide adequate security [3, 4]. According to Kaspersky Lab [5], any WPA-WPA3 encrypted Wi-Fi network is unprotected and can be attacked by reinstalling the key.

Since it is not possible to control the transmission space of information of a wireless network, then you need to focus on the control (identification) of users of such networks. Each device has its own unique features as a person fingerprints, retina, stroke or handwriting. The equipment of the device included in the Wi-Fi (frequency generator, modulator, radio transmitter, filter, antenna-feeder system) can have their own characteristics [6, 7, 8, 9].

Given the many shortcomings and vulnerabilities that allow malicious influences to successfully overcome the system of information protection, it is important to consider research aimed at comprehensive security, using additional parameters to detect unauthorized access and detection of intruders, namely the search for new methods of comparing spectra, that will allow their use in real protection systems wireless Wi-Fi networks.

**The object of** study is the process of recognizing similar spectra from different mobile devices from each other. This process is influenced by many factors: technical features of devices, position, distance, signal level and many others.

**The subject of study** is the evaluation of methods of comparing similar spectra by correlation analysis of the frequency dependence of the signal amplitude.

**The purpose of the work** is to recognize and identify similar spectrums of mobile devices of wireless Wi-Fi networks to increase the security indicators of the network.

## 1 PROBLEM STATEMENT

For the input parameters, which are: energy spectra $P_{L1}(f_i)$ ($i$ = 0, 1 ,…, $N$) (obtained experimentally [10], two of them are presented in Figure 1) $K$ legal network users, together with their MAC-addresses stored in the database in the form of a matrix with dimensions [$N$×$K$]; energy spectrum of signals of subscribers connected $P_{L2}(f_i)$ and subject to the identification procedure; the correlation function $B(j)$, is calculated, which is a function of the difference in frequency readings $j$ = –$N$+1… .$N$–1.

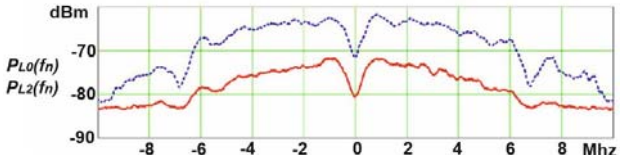


Figure 1 – Spectra of devices

The initial data are parameters $B(j)$ such as σ, Δ, $A$.

The task of the work consists in choosing such a parameter $B(j)$ (σ, Δ, A) that would allow in the future to quickly and with minimal error determine whether the subscriber who joined the wireless network is the one he claims to be (whose MAC address he uses). For the possibility of identification, the condition must be fulfilled: ACF (σ, Δ, A) → min, and for CCF (σ, Δ, A) → max.

## 2 REVIEW OF THE LITERATURE

The literature review showed that a fairly developed theory of calculating spectra by incomplete time series, using various window functions, various methods of averaging and smoothing them. In particular, the methods of Bartlett, Welch, Blackman-Tewkey, modified periodogramm and others. They can be used to detect weak frequency components in the signal, or bands where its maximum energy is concentrated, as well as for other similar tasks.

Many studies are conducted to identify by spectra, they use a variety of technologies and methods of comparison. From the calculation of the average amplitude value to neural networks [11, 12]. But all of them are aimed at distinguishing technologies according to frequency and modulation. In our case, we need already existing spectra, and almost identical ones that work at the same frequency, use the same modulation, accurately distinguish as fingerprints in humans.

The biometric identity perspective can be found in [13, 14].

In [15] the authors conduct research on the passive radiometric identification system of the PARADIS devices. This system is its own development of authors. The authors claim that PARADIS identifies unauthorized devices by comparing the radiometric signature of the device with known sanctioned signatures. PARADIS performs identification using two classifiers. One uses the support vector machine algorithm, while the other uses the $k$-nearest neighbor algorithm. Before a new network


DOI 10.15588/1607-3274-2024-4-1

card is allowed to access the network, the administrator measures and records its radiometric signature. The calculated signature is sent to the PARADIS server for identification. The system then compares the signature obtained with the known radiometric signatures of authorized network cards. If the radiometric signature does not match the network card that is allowed to use the secret key, PARADIS notifies the administrator of a possible security breach.

The accuracy of the studied systems is simply impressive – 99%. But the material in [15] does not provide detailed technical information on how radiometric signatures are calculated, it is not specified what exactly is understood by the concept of "radiometric signature" and on what it is based. And the nodes of the network are identified by the last four digits of the MAC address, which is the most vulnerable.

The non-standard identification approach for IoT is presented in [16]. This method is based on the use of not the entire spectrum, but the duration of the transition process, which is achieved in the obtained smooth version of the instantaneous amplitude characteristics of the signals of the transmitter using the method of averaging the sliding window.

The classification characteristics of the spectral prints proposed by the authors are evaluated using experimental data and described by a matrix of confusion. The efficiency of spectral prints is quantitatively determined by the criterion of class resolution. The effectiveness of the proposed method in the influence of noise through Monte Carlo simulation is shown.

In the article of the authors of this work [10] experimentally obtained spectra in visual analysis (with simple consideration of drawings) have something in common, and in some ways differ. The similarity of the spectra of Wi-Fi signals of the same device in different positions and differences in the spectra of radiation in different devices are established, which can be used to identify them. For comparison, the method of calculating the mean square of the difference between the corresponding spectral samples was used, taking into account the difference in the average power of different signals, which allows comparing the spectra obtained in different conditions with the patterns. The study [10] shows that each device has its own individual signal spectrum, which can be seen even visually and the results of analysis of spectra based on the calculation of the mean square of the difference indicate their difference. But the method under consideration gives in some cases similarity of results, so you need to look for other methods that can reduce the probability of a first kind error.

For signal analysis has a great practical application correlation statistical analysis of experimental data. The essence of correlation analysis is reduced to the establishment of the equation of regression (algebraic equation), that is, the type of rectilinear or curved relationship between the values, the estimation of the tightness (force) of the relations and the reliability of the measurement results [17]. Similar solutions have been found in the field of voice identification [18].

Therefore, the actual task is to find new methods for comparing the spectra of mobile devices, which will allow their use in real protection systems wireless Wi-Fi networks.

## 3 MATERIALS AND METHODS

To quantify the degree of difference between the $S(t)$ signal and its offset copy $S(t–\tau)$, the operation uses the $S(t)$ signal's ACF, which is equal to the dot product of the signal and its copy:

$$B(\tau) = \int_{-\infty}^{\infty} S(t)S(t-\tau)dt\,. \tag{1}$$

We apply the expression (1) to compare the spectra:

$$B(j) = \frac{1}{N}\sum_{i=0}^{N-1} P_{L1}(f_i)P_{L2}(f_{i+j})\,. \tag{2}$$

The specified expression (2) is used to calculate the correlation of the template (ACF) and the spectrum of the template with the spectra of "alien" devices (CCF).

The calculation of the MSD for the functions obtained is calculated by the formula:

$$\sigma = \sqrt{\frac{1}{2N\cdot B_{cp}}\sum_{j=-N+1}^{N-1}\left([(j)]^2\cdot B(j)\right)}, \tag{3}$$

$$B_{cp} = \frac{1}{2N}\sum_{j=-N+1}^{N-1} B(j)\,. \tag{4}$$

An important indicator of correlation processing is the asymmetry coefficient. The calculation of the asymmetry coefficient in the work is realized as the ratio of the central empirical moment of the third order to the cube of the mean squared deviation:

$$A = \frac{m_2}{\sigma^2}\,, \tag{5}$$

where, the central empirical moment of the third order was calculated by the formula:

$$m_2 = \frac{1}{2N\cdot B_{cp}}\sum_{j=-N+1}^{N-1}(j-j_{cp})^2 B(j)\,. \tag{6}$$

In relation to our case, asymmetry was calculated:

$$A = \frac{\frac{1}{2N\cdot B_{cp}}\sum_{j=-N+1}^{N-1}(j-j_{cp})^2 B(j)}{\sqrt{\frac{1}{2N\cdot B_{cp}}\sum_{j=-N+1}^{N-1}\left([(j)]^2\cdot B(j)\right)}}\,. \tag{7}$$


DOI 10.15588/1607-3274-2024-4-1

## 4 EXPERIMENTS

In accordance with the task, the spectra of mobile devices were calculated and modeled in the Mathcad environment. The simulation was conducted sequentially in accordance with (1)–(7).

The results of the simulation are displayed in the form of tables, graphs of dependence on the drug of ACF and CCF for the studied spectra and in the form of drawings.

## 5 RESULTS

Figure 2 shows that the peaks of correlation functions are at different levels, which makes it impossible to compare them. This is due to the fact that the signal level when removing the patterns of spectra of different mobile devices differ. In practice, the signal level from the devices may be quite insignificant. Therefore, for the convenience of comparing the correlation functions, they were normalized according to the maximum signal strength level when the peaks of the correlation functions are in 1, as shown in Figure 3.

There was no significant difference in the dependencies shown in Figure 3. The results of the calculation σ for one of the templates (A) in relation to the different provisions of the devices are shown in Table 1, showing similar results. Thus, this indicator can not be used to identify mobile devices.

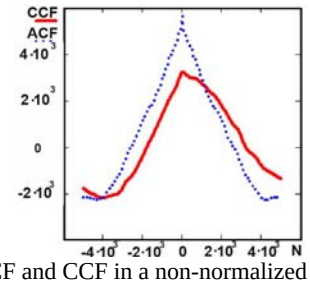



Figure 2 – ACF and CCF in a non-normalized form (ACF smartphone D with a smartphone D template)

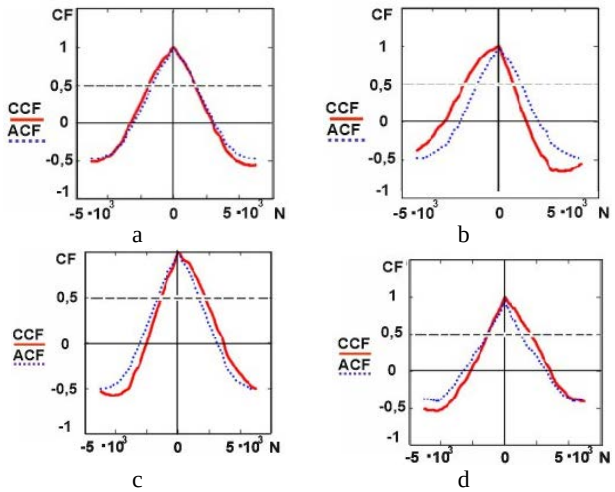



Figure 3 – ACF and CCF in the normalized form
A – ACF smartphone A with a smartphone A template;
B – ACF smartphone D with a smartphone template C;
B – ACF smartphone C with A smartphone A template;
G – ACF smartphone with a smartphone pattern D

Table 1 – σ calculations for pattern A

| template | A1 | A1 | A3 | A4 | B1 | B2 | B3 | B4 |
|---|---|---|---|---|---|---|---|---|
| σ | 945 | 943 | 940 | 937 | 937 | 949 | 952 | 952 |
| template | C1 | C2 | C3 | C4 | D1 | D2 | D3 | D4 |
| σ | 955 | 953 | 942 | 943 | 944 | 954 | 940 | 932 |

The study of the width (in the amount of reprint) of the ACF for templates and CCF for all devices with respect to the template at level 0.5 are shown in Table 2.

Table 2 – the results of measuring the width of the ACF and CCF by level 0.5

| Position | The width of the ACF (CCF) on level 0.5 | | | | | |
|---|---|---|---|---|---|---|
| | Pattern A | A | B | C | D | E |
| 1 | 2650 | 2655 | 2715 | 2864 | 2701 | 2929 |
| 2 | 2650 | 2688 | 2726 | 2879 | 2833 | 2937 |
| 3 | 2650 | 2675 | 2796 | 2817 | 2588 | 2844 |
| 4 | 2650 | 2635 | 2796 | 2785 | 2459 | 2807 |
| Middle | 2650 | 2663 | 2758 | 2836 | 2645 | 2879 |

From Table 2 it is clear that the width of the CCF of other devices can be narrower than the ACF.

On the basis of the completed calculations, we can say that:

– There is no significant difference in the mean squared deviation of the CCF for the template with its own device and other people's ones;

– The difference in the width of CCF by level 0.5 is also not detected. When normalizing, all functions are almost identical;

– A significant shift in the central frequency in the CCF is also not observed.

Thus, parameters up to the second order inclusive do not allow to detect the difference between the two spectra. But from Figure 3 it is clear that the CCF have a certain "bias", which can be characterized by an estimation of asymmetry for the empirical distribution.

Table 3 shows the results obtained for (7) the asymmetry coefficients. The average value of the asymmetry coefficient was calculated by summing each element by the value of the module. The number near the letter means the number of position of the mobile device, and the index "sr. average for four positions.

For greater observation, the results of this study are presented in graphic form in Figure 4, which shows the

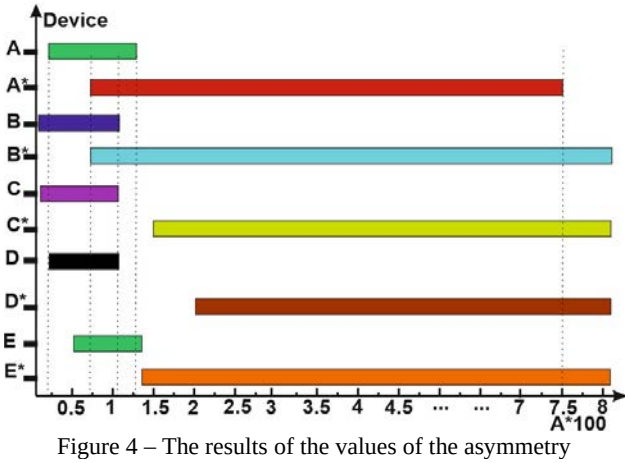



Figure 4 – The results of the values of the asymmetry coefficients in the graphic form


DOI 10.15588/1607-3274-2024-4-1

Table 3 – results of the value of the assymmetry coefficients

| | | Device patterns | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A | $A_{ср}$ | B | $B_{ср}$ | C | $C_{ср}$ | Д | $Д_{ср}$ | E | $E_{ср}$ |
| Devices in different positions | A1 | 0.013 | 0.0057 | 0.008 | 0.007 | –0.07 | 0.07 | 0.075 | 0.06 | 0.038 | 0.03 |
| | A2 | 0.002 | | –0.005 | | –0.07 | | 0.06 | | 0.03 | |
| | A3 | 0.004 | | –0.007 | | –0.07 | | 0.06 | | 0.029 | |
| | A4 | 0.004 | | –0.008 | | –0.066 | | 0.059 | | 0.017 | |
| | B1 | 0.005 | 0.008 | –0.007 | 0.0048 | –0.066 | 0.074 | 0.049 | 0.07 | 0.026 | 0.04 |
| | B2 | 0.003 | | –0.0004 | | –0.076 | | 0.074 | | 0.038 | |
| | B3 | 0.016 | | 0.011 | | –0.075 | | 0.087 | | 0.046 | |
| | B4 | 0.006 | | 0.0007 | | –0.079 | | 0.077 | | 0.036 | |
| | C1 | 0.077 | 0.07 | 0.079 | 0.07 | –0.008 | 0.0055 | 0.057 | 0.025 | 0.093 | 0.09 |
| | C2 | 0.077 | | 0.08 | | 0.011 | | –0.002 | | 0.065 | |
| | C3 | 0.063 | | 0.067 | | –0.002 | | 0.022 | | 0.093 | |
| | C4 | 0.064 | | 0.069 | | 0.001 | | 0.017 | | 0.094 | |
| | D1 | –0.077 | 0.06 | –0.083 | 0.07 | 0.038 | 0.04 | –0.023 | 0.01 | –0.08 | 0.06 |
| | D2 | –0.07 | | –0.075 | | –0.039 | | 0.016 | | –0.05 | |
| | D3 | –0.051 | | –0.057 | | –0.056 | | 0.012 | | –0.049 | |
| | D4 | –0.048 | | –0.022 | | –0.022 | | 0.002 | | –0.068 | |
| | E1 | –0.012 | 0.024 | –0.022 | 0.03 | –0,098 | 0.09 | 0.08 | 0.07 | 0.016 | 0.016 |
| | E2 | –0.012 | | –0.019 | | –0.114 | | 0.08 | | 0.022 | |
| | E3 | –0.038 | | –0.044 | | –0.078 | | 0.052 | | –0.014 | |
| | E4 | –0.038 | | –0.045 | | –0.078 | | 0.051 | | –0.012 | |

ranges of the obtained values of ACF and CCF. The letters in the figure show the ranges of values for the template and its device, and the letter with an asterisk shows the ranges of values for the template with other devices.

As can be seen from Figure 4, the values of the asymmetry coefficients for a template with its own device are in the range from 0 to 1.5. And the value for templates with other devices can be from 0.7 to 8. In most cases, the values of templates and other devices do not overlap. But still for some mobile tristroi (A and B) we see that there is a small range of values. In this range, additional research of the spectrum should be carried out (visually they differ significantly from each other) in order to prevent errors of the first kind.

## 6 DISCUSSION

The considered method of identification of devices in wireless networks Wi-Fi uses the sign of the state of the network at the physical level, which allows to detect and together with intrusion detection systems to prevent a number of attacks and thereby increase the security of Wi-Fi networks.

The obtained results showed the minimum value of the asymmetry coefficient when comparing the template with different positions of your own device.

Therefore, this feature can also be used to identify Wi-Fi devices in a model to detect abnormal network states [19]. Previous studies [10] which were based on the calculation of the mean square of the difference due to different signal levels can give a large error in the results, in contrast to the proposed method where the obtained values of the correlation functions were normalized.

Further research is advisable to develop in the direction of finding additional methods of analysis in ranges that do not give an unambiguous answer to the belonging of the mobile spectrum to the template. It is also necessary to confirm the effectiveness of this method in the presence of noise and experimental studies of the decision-making model on the abnormal state of the network [19] taking into account the parameter under consideration.

## CONCLUSIONS

1. The method of processing the results of measuring the radiation spectra of Wi-Fi devices by assessing the asymmetry coefficient of the intercorrelation function of the spectrum of Wi-Fi devices is proposed.

2. It is shown that the asymmetry coefficient of the CCF for the template and the corresponding device is significantly less than for the template and other device, which can serve as an identifying sign.


DOI 10.15588/1607-3274-2024-4-1

3. Set the range of the asymmetry coefficients, which can correspond to both the same device in different positions and different devices. To identify devices in this range, it is necessary to carry out a more detailed analysis of the spectrum.

## ACKNOWLEDGEMENTS

The authors express their gratitude to the Department of computer Radio Engineering and Systems of Technical Information Protection of Kharkiv National University of Radio Electronics for access to equipment for experimental research, the analysis and processing of which was written publication.

## REFERENCES

1. Sahabul A., Debashis D. Analysis of security theats in wireless sensor network, *International Journal of Wireless & Mobile Networks*, 2014, Vol. 6, № 2, pp. 35–46. DOI: 10.5121/ijwmn.2014.6204
2. Gupta A., Jha R. K. Security threats of wireless networks: A survey, *International Conference on Computing, Communication & Automation, Greater Noida, 15–16 May 2015: proceedings*. Greater Noida, IEEE, 2015, pp. 389–395. DOI: 10.1109/CCAA.2015.7148407
3. Catania. C. A., Garino C. G. Automatic network intrusion detection: Current techniques and open issues, *Computers and Electrical Engineering*, 2012, Vol. 38, pp. 1062–1072. DOI: 10.1016/j.compeleceng.2012.05.013
4. Garcıa-Teodoroa P., Dıaz-Verdejoa J., Macia´-Fernandeza G. et al. Anomaly-based network intrusion detection: Techniques, systems and challenges, *Computers & Security*, 2009, Vol. 28, pp. 18–28. DOI: 10.1016/j.cose.2008.08.003
5. “Your Wi-Fi connection is unsecured” warning [Electronic resource]. Access mode: https://support.kaspersky.com/common/macos/14582
6. Agilent 802.11a/g Manufacturing Test Application Note : A Guide toGetting Started. Application note 1308-3 [Electronic resource]. Access mode: https://studylib.net/doc/18797817/agilent-802.11a-g-manufacturing-test-application-note-—-a...
7. Making 802.11g transmitter measurements. Application note 1380-4 [Electronic resource]. Access mode: https://www.testunlimited.com/pdf/an/5988-7813EN.pdf
8. RF Testing of WLAN products. Application note 1380-1 [Electronic resource]. Access mode: https://www.testunlimited.com/pdf/an/5988-3762EN.pdf
9. Agilent Technologies. Testing and Troubleshooting Digital RF Communications Transmitter Designs. Application note 1313 [Electronic resource]. Access mode: https://archive.org/details/manualzilla-id-6880267/page/12/mode/2up
10. Antipov I. Ye., Vasylenko T. O. Ydentyfykatsyia mobylnykh ustroistv po osobennostiam spektrov ykh syhnalov, *Radiotekhnika. Vseukr. mizhvid. nauk.-tekhn. zb.*, 2020, Vyp. 201, pp. 91–97.
11. Perez-Neira A. I., Member S., Lagunas M. A., Rojas M. A., Stoica P. Correlation Matching Approach for Spectrum Sensing in Open Spectrum Communications [Electronic resource]. Access mode: https://www.academia.edu/7471893/Correlation_matching_approach_for_spectrum_sensing_in_open_spectrum_communications
12. Tekbıyık K., Akbunar Ö., Ekti A. R.et al. Correlation matching approach for spectrum sensing in open spectrum communications, *IEEE Transactions on Signal Processing*, 2020, Vol. 57, № 12, pp. 18–28. DOI: 10.1109/TSP.2009.2027778
13. Ross A., Jain A. Information fusion in biometrics, *Pattern Recognition Letters*, 2003, Vol. 24, pp. 2115–2125. DOI: DOI:10.1016/S0167-8655(03)00079-5
14. Tuyls P., Goseling J. Capacity and Examples of Template-Protecting Biometric Authentication Systems, *Biometric Authentication, ECCV International Workshop, Prague, 15 May, 2004: proceedings*. Prague, BioAW, 2004. pp. 1–13. DOI:10.1007/978-3-540-25976-3_15
15. Brik V., Banerjee S., Gruteser M. et al. Wireless device identification with radiometric signatures, *MobiCom ‘08: Proceedings of the 14th ACM international conference on Mobile computing and networking, San Francisco 14–19 September 2008: proceedings*. San Francisco, 2008, pp. 116–127. DOI:10.1145/1409944.1409959
16. Köse M. Taşcioğlu S., Telatar Z. RF Fingerprinting of IoT Devices Based on Transient Energy Spectrum, *IEEE Access*, 2019, Vol. 7, pp. 18715–18726. DOI: 10.1109/ACCESS.2019.2896696
17. Sydenko V. M., Hrushko Y. M. Osnovy nauchnykh yssledovanyi. Kharkov, Vyshcha shkola, 1978, 200 p.
18. Craciun A., Gabrea M. Correlation coefficient-based voice activity detector algorithm, *Canadian Conference on Electrical and Computer Engineering, Canada, 2–5 May, 2004: proceedings*. Niagara Falls, ON, Canada, IEEE, 2004, pp. 1789–1792. DOI: 10.1109/CCECE.2004.1349763
19. Antipov I., Vasilenko T. Improving the model of decision making about abnormal network state using a positioning system, *Eastern-European Journal of Enterprise Technologies*, 2019, Vol. 1, № 9 (97), pp. 6–11. DOI: 10.15587/1729-4061.2019.157001

Received 06.05.2024.
Accepted 26.09.2024.


DOI 10.15588/1607-3274-2024-4-1

УДК 621.396.946

# ІДЕНТИФІКАЦІЯ МОБІЛЬНИХ ПРИСТРОЇВ ЗА КОРЕЛЯЦІЙНИМИ ОСОБЛИВОСТЯМИ ЇХ СИГНАЛІВ

**Антіпов І. Є. –** д-р техн. наук, професор, завідувач кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Харківський національний університет радіоелектроніки, м. Харків, Україна.

**Василенко Т. О.** – канд. техн. наук, старший викладач кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Харківський національний університет радіоелектроніки, м. Харків, Україна.

## АНОТАЦІЯ

**Актуальність.** Масовому поширенню Wi-Fi мереж сприяє простота їх розгортання, висока швидкість, універсальність і зручність використання. Розвиток і поширення цих мереж триває, незважаючи на наявність ряду недоліків. Одним з недоліків є їх вразливість до різних видів атак, у тому числі, основаних на підробці (імітації) ідентифікаційних даних. Разом з тим існують ознаки фізичного рівня, знання яких розширює уявлення про стан мережі, може сприяти підвищенню надійності ідентифікації абонентів мережі і таким чином запобіганню ряду атак. Це дослідження направлене на теоретичне обґрунтування можливості їх застосування.

**Мета.** Метою дослідження є оцінка застосування детального аналізу спектрів сигналів, випромінюваних пристроями, підключеними до безпровідних мереж Wi-Fi, для їх ідентифікації. Для досягнення поставленої мети необхідно на основі експериментально виміряних спектрів безпровідних пристроїв, підключених до мережі Wi-Fi, провести аналіз отриманих результатів та оцінити можливість використання спектра для ідентифікації мобільних пристроїв.

**Метод.** В даній роботі запропоновано метод обробки результатів вимірювання спектрів випромінювання Wi-Fi пристроїв шляхом оцінки коефіцієнта асиметрії взаємокореляційної функції спектру Wi-Fi пристроїв. Для оцінки ефективності методу використовувалося математичне моделювання.

**Результати.** Результати досліджень показують, що мінімальне значення коефіцієнта асиметрії при порівнянні шаблону з різними положеннями власного пристрою, і великі значення коефіцієнта асиметрії при порівнянні шаблонів з чужими спектрами. Отже, ця ознака також може бути використана для ідентифікації Wi-Fi пристроїв.

**Висновки.** Результати досліджень говорять про можливість застосування запропонованого методу для ідентифікації мобільних пристроїв , що дозволить якісно доповнити існуючі моделі забезпечення безпеки ще однією ознакою виявлення несанкціонованого доступу.

**КЛЮЧОВІ СЛОВА:** безпека, Wi-Fi, ідентифікація, спектр, коефіцієнт асиметрії, мобільній пристрій.

## ЛІТЕРАТУРА

1. Sahabul A. Analysis of security theats in wireless sensor network / A. Sahabul, D. Debashis // International Journal of Wireless & Mobile Networks – 2014. – Vol. 6, № 2. – P. 35–46. DOI: 10.5121/ijwmn.2014.6204
2. Gupta A. Security threats of wireless networks: A survey / A. Gupta, R. K. Jha // International Conference on Computing, Communication & Automation, Greater Noida, 15–16 May 2015: proceedings. – Greater Noida: IEEE, 2015. – P. 389–395. DOI: 10.1109/CCAA.2015.7148407
3. Catania. C. A. Automatic network intrusion detection: Current techniques and open issues / C.A. Catania, C. G. Garino // Computers and Electrical Engineering. – 2012. – Vol. 38. – P. 1062–1072. DOI: 10.1016/j.compeleceng.2012.05.013
4. Anomaly-based network intrusion detection: Techniques, systems and challenges / [P. Garcıa-Teodoroa, J. Dıaz-Verdejoa, G. Macia´-Fernandeza et al.] // Computers & Security. – 2009. – Vol. 28. – P. 18–28. DOI: 10.1016/j.cose.2008.08.003
5. "Your Wi-Fi connection is unsecured" warning [Electronic resource]. – Access mode: https://support.kaspersky.com/common/macos/14582
6. Agilent 802.11a/g Manufacturing Test Application Note : A Guide toGetting Started. Application note 1308-3 [Electronic resource]. – Access mode: https://studylib.net/doc/18797817/agilent-802.11a-g-manufacturing-test-application-note-—-a...
7. Making 802.11g transmitter measurements. Application note 1380-4 [Electronic resource]. – Access mode: https://www.testunlimited.com/pdf/an/5988-7813EN.pdf
8. RF Testing of WLAN products. Application note 1380-1 [Electronic resource]. – Access mode: https://www.testunlimited.com/pdf/an/5988-3762EN.pdf
9. Agilent Technologies. Testing and Troubleshooting Digital RF Communications Transmitter Designs. Application note 1313 [Electronic resource]. – Access mode: https://archive.org/details/manualzilla-id-6880267/page/12/mode/2up
10. Антіпов І. Є. Идентификация мобильных устройств по особенностям спектров их сигналов / І. Є. Антіпов, Т. О. Василенко // Радіотехніка. Всеукр. міжвід. наук.-техн. зб. – 2020. – Вип. 201. – С. 91–97.
11. Correlation Matching Approach for Spectrum Sensing in Open Spectrum Communications / [A. I. Pérez-Neira, S. Member, M. A. Lagunas et al.] [Electronic resource]. – Access mode: https://www.academia.edu/7471893/Correlation_matching_approach_for_spectrum_sensing_in_open_spectrum_communications
12. Correlation matching approach for spectrum sensing in open spectrum communications [K. Tekbıyık, Ö. Akbunar, A. R. Ekti et al.] // IEEE Transactions on Signal Processing. – 2020. – Vol. 57, № 12. – P. 18–28. DOI: 10.1109/TSP.2009.2027778
13. Ross A. Information fusion in biometrics / A. Ross, A. Jain // Pattern Recognition Letters. – 2003. – Vol. 24. – P. 2115–2125. DOI: DOI:10.1016/S0167-8655(03)00079-5
14. Tuyls P. Capacity and Examples of Template-Protecting Biometric Authentication Systems / P. Tuyls, J. Goseling // Biometric Authentication, ECCV International Workshop, Prague, 15 May, 2004: proceedings. – Prague : BioAW, 2004. – P. 1–13. DOI:10.1007/978-3-540-25976-3_15
15. Wireless device identification with radiometric signatures [V. Brik, S. Banerjee, M. Gruteser et al.] MobiCom '08: Proceedings of the 14th ACM international conference on Mobile computing and networking, San Francisco 14–19 September 2008: proceedings. – San Francisco, 2008. – P. 116–127. DOI:10.1145/1409944.1409959
16. Köse M. RF Fingerprinting of IoT Devices Based on Transient Energy Spectrum / M. Köse, S. Taşcioğlu, Z. Telatar // IEEE Access. – 2019. – Vol. 7. – P. 18715–18726. DOI: 10.1109/ACCESS.2019.2896696
17. Сиденко В. М. Основы научных исследований / В. М. Сиденко, И. М. Грушко. – Харьков : Выща школа, 1978. – 200 с.
18. Craciun A. Correlation coefficient-based voice activity detector algorithm / A. Craciun, M. Gabrea // Canadian Conference on Electrical and Computer Engineering, Canada, 2–5 May, 2004: proceedings. – Niagara Falls, ON, Canada, IEEE, 2004. – P. 1789–1792. DOI: 10.1109/CCECE.2004.1349763
19. Antipov I. Improving the model of decision making about abnormal network state using a positioning system / I. Antipov, T. Vasilenko // Eastern-European Journal of Enterprise Technologies. – 2019. – Vol. 1, № 9 (97). – P. 6–11. DOI: 10.15587/1729-4061.2019.157001


DOI 10.15588/1607-3274-2024-4-1

UDC 681.7.068

# METHOD OF CONTROL THE MECHANICAL STATE OF THE OPTICAL FIBER OF THE DIELECTRIC SELF-SUPPORTING OPTICAL CABLE DURING OPERATION

**Bondarenko O. V.** – Dr. Sc., Professor, Laureate of the State Prize in Science and Technology, Odesa, Ukraine.
**Stepanov D. M.** – PhD, Associate Professor, Associate Professor of the Department of Switching Systems of Electronic Communications, State University of Intelligent Technologies and Telecommunications, Odesa, Ukraine.

## ABSTRACT

**Context.** One of the issues of theoretical and practical research studying phenomena that occur over time and lead to violations of the normative work of optical cables (OC) are ways to ensure and control their reliability during operation.

Today, electronic communication (telecommunications) has already gained significant integration and widespread use due to the urgent need to exchange large volumes of information between users or network devices at high speeds and over long distances, as well as the provision of a wide range of electronic communication services.

The electronic communication service has a high level of demand and consists in receiving and/or transmitting information through electronic communication networks, which is transmitted using electronic communication networks and services.

In an electronic communication network, the transmitting/receiving of optical signals is provided by the fiber optic transmission system (FOTS). It is capable of converting electrical signals from a variety of digital devices into optical signals and transmitting them over fiber-optic communication lines (FOCL), which is the main transmission medium in an electronic communication network.

The problem of ensuring the reliability of the FOCL, which includes a wide range of issues related to the development and production of all its elements, design, construction and technical operation of the communication line, continues to gain more and more importance.

In general, the transmitting/receiving of information between end users equipment, communication nodes, network devices (servers, databases, etc.) takes place through an electronic communication network.

Normative and technical documentations for fiber-optic communication lines regulates the control of the mechanical state of the optical cable during operation, but do not provide the full control of the mechanical state of the optical fiber to ensure the quality and reliability of the line during the specified service life.

As known, to ensure the reliability of the optical cable, as a rule, the permissible elongation of the optical fiber (OF) is $\varepsilon_{pOF} < (0.2...0.25)$ %, adopted during the designing of the cable. However, during operation, the appearance of multiple excess of elongations exceeding these values is possible in the fibers.

Thus, the development and substantiation of methods for evaluating the mechanical characteristics of a dielectric self-supporting optical cable (DSOC) and the method of full control of the mechanical state of the optical fiber is necessary. The last can lead to premature failure of the optical fiber.

**Objective.** Development and substantiation of the method of control the mechanical state of an optical fiber of the suspended DSOC, as well as assessment of the conditions of deformation of optical fibers in its core with the appearance of longitudinal tensile/compressive loads during operation.

**Method.** Two ways of evaluating the mechanical characteristics of DSOC and the method of control the mechanical state of its fibers have been developed and proposed. For this, the following characteristics of the cable and fiber are adopted in the work: relative elongation of the cable and fiber ($\varepsilon_{cx}$, $\varepsilon_{OFx}$), span length ($L_{sx}$) of the line, cable sag in the span ($f_x$) and tensile load (TL) of the cable ($F_{tlx}$), which causes longitudinal deformation ε. At the same time, the method proposes to control the mechanical state of the optical fiber during the operation of the DSOC by determining its effective relative elongation according to the mechanical, physical and climatic conditions of the line location.

In the paper, it is proposed to measure and calculate the following mechanical characteristics, due to the developed reference data for the selection of the cable type and the climatic zone of the line location, measuring equipment and mathematical tools:

– equivalent mechanical tension in DSOC;
– calculated and actual cable sag in the span;
– actual effective relative elongation of the cable;
– actual tensile load acting on the cable.

The ways and method presented in the work allow a complete evaluation of the mechanical characteristics of the cable and control of the mechanical state of the optical fiber during operation of the DSOC. It creates an opportunity to monitor its changes to prevent the appearance of excessive loads during operation and failure of the fiber-optic communication line.

It is possible to recommend this method for use by relevant departments for technical operation of telecommunication lines and networks based on hanging optical cables.

**Results.** The work presents the results of the development and justification of the method of control the mechanical state of optical fibers of dielectric self-supporting optical cables during operation. For example, using the developed method, it is shown that in the cable OKL-3-D2A14 produced by PJSC “Odeskabel” in the conditions of the Odesa climate zone (Black Sea region), optical fibers with a span length of 100 m are subject to elongation by 0.16 %, and DSOC is subject to an actual tensile force of 2.722 kN. This result of the control of the mechanical state of the OF established that such span of the line ensures its mechanical integrity within the limits of the permissible deformation of 0.25 % adopted in the design of the cable, but exceeds its permissible tensile load of 2.6 kN.

**Conclusions.** The scientific novelty of the work results is that, for the first time ways of fully evaluating the mechanical characteristics of the DSOC during operation and the method of fully control the mechanical state of its optical fiber have been developed. It allows to monitoring changes in the mechanical state of the optical fiber of the cable.

**KEYWORDS:** relative elongation, dielectric self-supporting optical cable, optical fibers, mechanical stresses, longitudinal tensile load, physical and climatic loads.


DOI 10.15588/1607-3274-2024-4-2

## ABBREVIATIONS

FOCL is a fiber-optic communication line;

OC is an optical cable;

OF is an optical fiber;

DSOC is a dielectric self-supporting optical cable;

TL is a tensile load;

OFOCL is an overhead fiber-optic communication line;

CSE is a central strength element;

PSE is a peripheral strength element;

RTD is a domestic regulatory and technical documentations;

ADSS is an All-Dielectric Self-Supporting cable;

BR is a Brillouin reflectometer;

MPTL is a maximum permissible tensile load;

FR is a fiberglass rods;

AT is an aramid thread;

MPC is a maximum permissible relative elongation of a cable;

OM is an optical module tube of a cable.

## NOMENCLATURE

$\varepsilon_{pOF}$ is a permissible elongation of the optical fiber;

$\varepsilon_{exOF}$ is an excess length of the optical fiber;

$\varepsilon_{pcx}$ is a permissible elongation of the optical cable at some condition;

$\varepsilon_{cx}$ is a relative elongation of the cable at some condition;

$\varepsilon_{apcx}$ is an actual additional relative elongation of the cable under certain climatic conditions;

$\varepsilon_{mpc}$ is a maximum permissible relative elongation of a cable;

$\varepsilon_{OFx}$ is a relative elongation of the fiber at some condition;

$L_{sx}$ is a span length of the line at some condition;

$f_x$ is a cable sag arrow in the span at some condition;

$F_{tlx}$ is a tensile load of the cable at some condition;

$F_{mtl}$ is a maximum permissible tensile load;

$\varepsilon$ is a longitudinal elongation;

$P_c$ is a weight of the cable;

$D_c$ is a diameter of the cable;

$\alpha_{OC}$ is a temperature coefficient of linear expansion of the cable;

$E_{eq}$ is an Young's modulus of the cable;

$\upsilon$ is a wind pressure;

$\Delta t_h$ is a thickness of hoarfrost;

$\Delta t_i$ is a thickness of ice;

$\sigma_1$ is an equivalent mechanical stress of the cable;

$L_s$ is a length of the span of the OFOCL;

$\gamma_1$ is a specific load of the own weight of the cable;

$f$ is a cable sag arrow in the span;

$S$ is a cross-sectional area of the cable;

$n$ is an overload factor;

$\sigma_x$ is a equivalent mechanical stress;

$\gamma_x$ is a specific load on the cable under physical and mechanical conditions of the climatic zone;

$t_x$ is a cable temperature under certain conditions;

$S_{SE}$ is a cross-sectional area of all strength elements of the cable;

$g$ is a Galilean constant;

$h$ is a dimensions of the OFOCL in the span;

$R$ is a radius of the spiral arrangement of the element around the CSE;

$\Delta R$ is a distance between OF (or OF bundle) and the inner surface of the wall of OM tube;

$h_s$ is a step of spiral laying of elements (OM or filler element) of the cable core around the CSE;

$\sigma_{ax}$ is an actual mechanical stress of the cable along the length of the line span;

$f_{ax}$ is an actual cable arrow sag in the span of the line;

$F_{atlx}$ is an actual tensile load of the cable under the influence of certain conditions of the operating environment;

$\sigma_{ax}$ is an actual mechanical stress of the cable obtained by measuring $f_{ax}$;

$\sigma_{axc}$ is an actual mechanical stress of the cable obtained by calculating $f_{axc}$;

$f_{ax}$ is a measured arrow of cable sag in the span of the line;

$f_{axc}$ is a calculated arrow of cable sag in the span of the line.

## INTRODUCTION

During the operation of a dielectric self-supporting optical cable, its optical fiber, as a rule, is under certain local or distributed along its length mechanical loads that create mechanical stresses. The mechanical stresses of the DSOC itself, suspended on the supports of the overhead fiber-optic communication line (OFOCL), appear under the influence of mechanical (cable weight, the radius of its bending in the clamps of the pole) and physical and climatic factors of the operating area (temperature, wind pressure, hoarfrost, ice, etc.). As known, changing the longitudinal mechanical actions on the cable leads to its tension or compression, that is, changes in its span length, sag and tensile load. Therefore, when choosing the design of the DSOC and calculating the conditions for its suspension and operation, it is necessary to take into account the constant action on the cable of mechanical, physical and climatic factors of the location of the OFOCL. At the same time, ensuring the proper technical condition of the line during the expected service life requires monitoring, first of all, the mechanical characteristics of the cable.

In turn, the sag of the DSOC in the span of the line under the worst conditions of its operation makes it possible to check the actual tension in the cable according to the value of the initial tension, which was during the building of the line. Therefore, during the year, the operation of the OFOCL takes place with a change in the mechanical condition of the DSOC and its fibers.

A complete assessment of the mechanical condition of the DSOC is based on the determination of the specific loads on the cable in the span of the OFOCL and its mechanical characteristics. The action of excess stresses in the optical fiber leads to the appearance of multi-zonal cracks in its material, which eventually destroy it, thereby reducing its service life. In general, ensuring the quality and reliability of the line during the specified period of


DOI 10.15588/1607-3274-2024-4-2

operation requires monitoring the mechanical characteristics of the cable and the mechanical state of the fiber in combination with its maintenance.

Control of the mechanical state of the optical fiber during cable monitoring should be simple and easily accessible to the line operator's linear personnel.

In the regulatory and technical documentations for the projecting, building and operation of the OFOCL, there are no ways for assessing the mechanical characteristics of the DSOC and the method of control the mechanical state of its optical fiber. Therefore, the development of a method for improving the control of the mechanical state of optical fibers of dielectric self-supporting optical cables along the length of the OFOCL span is necessary and timely.

**The object of study** is the control of the mechanical state of the optical fiber of a dielectric self-supporting optical cable with a central strength element (CSE) made of a fiberglass rod and a peripheral strength element (PSE) made of aramid threads during the operation of the DSOC of the electronic communication network.

**The subject of study** is the mechanical state of the optical fiber of the DSOC, depending on the length of the OFOCL span and the conditions of its location in the electronic communication network.

**The purpose of the work** is development and justification:

– ways of assessing the mechanical characteristics of the DSOC along the length of the OFOCL span during operation;

– the method of control the mechanical state of the optical fiber of cable;

– analysis and discussion of the results of research of DSOC and its fiber according to the developed ways and method.

## 1 PROBLEM STATEMENT

Suppose there is DSOC model with a modular design, which contains a certain number of optical fibers and has technical characteristics that depend on its design features and purpose. First of all, these are: weight $P_c$ and diameter $D_c$ of the cable, equivalent temperature coefficient of linear expansion of the cable $\alpha_{OC}$ and the Young's modulus of the cable $E_{eq}$, permissible relative elongation of the cable $\varepsilon_{cx}$ and fiber $\varepsilon_{OFx}$, maximum permissible tensile load of the cable $F_{tlx}$. Such DSOC will be suspended on the poles of the OFOCL in a certain climatic zone of operation, which is characterized by the length of the line span $L_{sx}$, the cable sag arrow in the span $f_x$, the range of environmental temperatures, the pressure of the wind $\upsilon$, the thickness of hoarfrost $\Delta t_h$ and ice $\Delta t_i$ at the worst climatic conditions. Under the influence of mechanical, physical and climatic loads, excessive loads may appear in the structure of the cable and optical fiber, which can lead to fiber breakage and failure of the communication line.

The task of control the mechanical state of the optical fiber and cable, determining the suitability of DSOC for successful operation in such climatic zone consists in determining and monitoring the implementation of the following conditions: $\varepsilon_{OFx} < \varepsilon_{pOF}$ , $F_{tlx} \le (0.65...0.70)F_{mtl}$.

In turn, the problem of assessing the mechanical state of the optical fiber and DSOC consists in that it is necessary to determine the actually acting loads on the OF and DSOC, which depend on the physical and climatic conditions of OFOCL operation, i.e. $\varepsilon_{OFx}$, $F_{tlx} = f\,(P_c, D_c, \alpha_{OC}, E_{eq}, L_{sx}, f_x, t_x, \upsilon, \Delta t_h, \Delta t_i)$.

For this purpose, it is necessary:

– to create a reference base of the method with the geometric and structural parameters of the DSOC elements, equivalent temperature coefficient of linear expansion $\alpha_{OC}$ and Young's modulus $E_{eq}$ of the cable, physical and climatic conditions of operation;

– to develop and justify ways of assessing the mechanical characteristics of the DSOC during operation and the method of control the mechanical state of its optical fiber;

– analyze the results of control the mechanical characteristics of the DSOC along the length of the line span and, based on the values of the relative elongation of the optical cable, give conclusions about the mechanical state in the selected climatic zone;

– to determine the value of the actual mechanical characteristics of the DSOC and the mechanical state of its optical fiber;

– analyze the results of control the actual mechanical state of the optical fiber.

## 2 REVIEW OF THE LITERATURE

Analysis of international and domestic regulatory and technical documentations (RTD) and number of works by well-known authors in the field of fiber-optic communications, in particular, the development, production and operation of optical cables [1 – 12] showed the lack of completeness of research in this direction, and especially in the field of control of mechanical state of an optical fiber in DSOC during operation.

Thus, in the IEEE international standard "Standard for Testing and Performance for All-Dielectric Self-Supporting (ADSS) Fiber Optic Cable for Use on Electric Utility Power Lines" [1] there are requirements for the design, mechanical, electrical and optical characteristics of DSOC used on suspended engineering networks. Special attention is paid to maintaining proper optical fiber reliability and assessing overhead power line sag using a simple optomechanical system with optical fiber parameters.

In [2] it is shown approaches and recommendations for the selection of materials and the calculation of modular structures of cables and, first of all, strength elements in accordance with the requirements of international standards. The authors of the paper [3] proposed the method of evaluating the economic efficiency of multi-module optical cable designs based on mass-dimensional indicators, the results of which can be used for the selection of DSOC when designing OFOCL.

Some foreign and domestic authors [4–6] paid some attention to the study of some of the mechanical charac-


DOI 10.15588/1607-3274-2024-4-2

teristics of DSOC and control of the mechanical state of the optical fiber. As it is known, during the projecting, building and operation of the OFOCL, the assessment of the mechanical characteristics of the DSOC and the control of the mechanical state of the optical fiber are necessary. However, the results of studies in [4, 5] of the stresses of the DSOC and the mechanical state of the optical fiber under the influence of tensile forces and the operating temperature of the line are insufficient. The last is proof of the need for a full assessment of the mechanical characteristics of DSOC and control of the mechanical state of the optical fiber.

Real elongation of an optical fiber can be both short-term and long-term. It can reach an undesirable value during cable operation, that is, greater than the calculated value in the cable design [5].

In [6], the results of the study of the influence of mechanical stress in a special single-mode optical fiber, which can maintain the state of polarization of transmitted light and have the ability to resist the effects of the environment, are presented. These results also confirm the need to control the mechanical state of the optical fiber of the DSOC during the operation of the OFOCL.

In addition, the current regulatory and technical documentations of Ukraine on the projecting, building and operation of OFOCL provides for compliance with the required technical state of the cable and its fiber. In the rules for the arrangement of electrical installations [7], it is recommended to perform the calculation of the mechanical state of the DSOC based on the value of the design loads using the allowable stress method, taking into account the residual deformation in the cable and the allowable loads on the fiber.

The regulatory and technical document of Ukraine in recommendations for hanging optical cables P 45-010-2002 [8] provides recommendations for hanging optical cables on the supports of overhead lines, power lines, railway contact networks and recommends cables with maximum tensile and crushing loads that have an increase in the coefficient OF attenuation is no more than 0.05 dB. The last also requires monitoring of the mechanical state of the DSOC and its fibers during the operation of the OFOCL.

In [9, 10], the method of control the reliability of an optical cable using a Brillouin reflectometer (BR) is shown, which works on the basis of the method of reproducing the physics of the effect of multi-zone cracks in the quartz fiber material. Depending on the stress in the material and the area of its location, the reliability and service life of the OFOCL is evaluated. In general, by BR establishes the nature of the dependence of the service life of an optical fiber on its tension, namely on the action of stresses.

The main drawback of the method is the complexity of operating the reflectometer and its cost, which is beyond the reach of most operators of fiber-optic communication lines.

Thus, the review of literatures sources shown that in a limited number of works the requirements and methods of assessment and even monitoring of some mechanical characteristics of the optical cable in the span of OFOCL are given. But they do not give a full assessment of the mechanical state of the DSOC and its optical fiber. In this regard, the development and substantiation of a cheap and easily accessible method of controlling the mechanical state of OF in DSOC during operation is necessary.

## 3 MATERIALS AND METHODS

To begin with, it is necessary to perform theoretical foundations of the method.

The peculiarity of the operation of DSOC suspended on the poles of OFOCL, compared to the operation of cables laid, for example, in the soil, in cable ducts, etc., is the effect on them of dynamic mechanical, physical and climatic loads. As noted earlier, these loads cause a change in the length of the cable in the span between the line supports and the appearance of its mechanical tensile (compression) of stresses. The last makes it necessary to determine the cable sag in the span and the tensile load to which it corresponds.

The conditions for hanging the DSOC and its mechanical state during operation are determined by the values of: cable mechanical stress ($\sigma_{cx}$), relative elongations of the cable and its optical fiber ($\varepsilon_{cx}$, $\varepsilon_{OFx}$), span length ($L_{sx}$), sag ($f_x$) and tensile load (TL) of the cable ($F_{tl}$), which is caused by the longitudinal deformation $\varepsilon$. At the same time, under any operating conditions, it is necessary to ensure that the tensile force of the cable $F_{tlx}$ is less than its maximum permissible tensile load $F_{mtl}$ (MPTL), which is indicated in the technical conditions, i.e. $F_{tlx} \le (0.65...0.70) \cdot F_{mtl}$. In turn, the tensile loads acting on the structure of the DSOC will make it possible to determine the mechanical state of the OF in the core of the cable and its changes.

According to, for example, [11] hanging and installation of communication network cables can be carried out at temperatures from minus 5 ºC to plus 50 ºC, and operation – from minus 40 ºC to plus 70 ºC. In addition, different designs of DSOC provide different values of the maximum allowable tensile loads due to central and peripheral strength elements or their combination. This is due to the difference in the conditions of their suspension and operation (length of span, climatic conditions, etc.).

In the absence or effect of external physical and climatic loads (hoarfrost weight, ice weight) on the cable, in addition to when it is affected by its own weight and the temperature of the environment, it is possible to evaluate the mechanical state of the DSOC structure by the of sag arrow in the line span during operation. In order to control the mechanical state of the optical fiber, it is necessary to perform calculations of the sag arrow $f_x$ for an extended temperature range (–5...+70) ºC. It is possible to estimate the results of the $f_x$ calculations based on the data of its measurements using special equipment. In addition, the measured values of $f_x$ make it possible to find out about the value of the actual mechanical stress $\sigma_{ax}$ (or longitudinal deformation $\varepsilon_{ax}$) in it and to determine the actual tensile load acting at this moment. This will make it possible


DOI 10.15588/1607-3274-2024-4-2

to monitor changes in the mechanical characteristics of the cable over time to prevent the appearance of excessive TL and breakage of optical fibers.

In order to solve the given problem, it is possible to choose the following modular constructions of DSOC as a research subject, which contain:

– CSE and PSE from fiberglass rods (FR);

– CSE from fiberglass rods and PSE from winding aramid threads (AT).

Since the ways of assessing the mechanical state of these cable structures are the same, and the results will differ only due to the values of their equivalent temperature coefficient of linear expansion and Young's modulus of the cable, the second structure of the DSOC was chosen for research.

At the same time, the following assumptions are accepted in the work:

– fiberglass rods work within the limits of elastic deformation;

– the relative longitudinal elongation of aramid threads creep is not taken into account in the calculation of the PSE tensile load;

– the cable suspension points on the overhead line are at the same height.

As an example, let's solve the given problem for a cable produced by PJSC "Odeskabel" (Fig. 1), which is suspended on a OFOCL located in the conditions of the Odesa climate zone.

In this work, the following expressions are used to estimate the specific loads of the DSOC and its mechanical characteristics during operation.

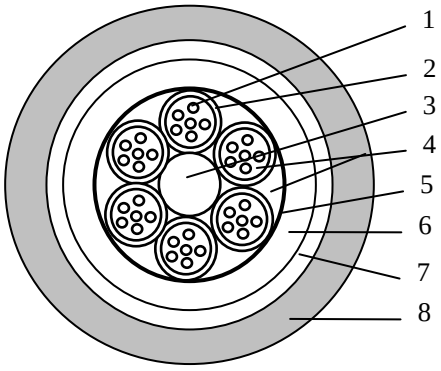



Figure 1 – Design of DSOC with strength elements from FR and AT: 1 – optical fiber; 2 – optical module tube (OM); 3 – CSE (fiberglass rod); 4 – filling compounds; 5 – fastening element; 6 – intermediate shell; 7 – PSE (winding of aramid threads); 8 – outer shell

The equivalent mechanical stress $\sigma_1$, which appears in the cable due to its suspension and the action only the self-weight of the DSOC, can be determined using the transformed formula for determining the sag arrow [11]:

$$\sigma_1 = \frac{L_s^2 \gamma_1}{8f} . \quad (1)$$

The last expression is valid with sufficient accuracy for engineering calculations when the cable is suspended at the same heights.

The specific loads on the DSOC ($\gamma_1 ... \gamma_7$) characterize the physical and climatic loads on the cable under operating conditions and include the load from its own weight, wind pressure, the weight of hoarfrost and ice.

The specific load from the self-weight of the DSOC can be determined by the expression [11]:

$$\gamma_1 = \frac{P_C}{S} n . \quad (2)$$

It is possible to determine the equivalent mechanical stress in the DSOC $\sigma_x$ under the influence of physical and climatic loads using the equation of state of the suspended cable in the span under the condition of the equality of the suspension points (Fig. 2) [11]:

$$A\sigma_x^3 + B\sigma_x^2 + C = 0 . \quad (3)$$

here $A = \left(1 + \alpha_{OC}(t_x - t)\right)\left(1 + \frac{L_s^2 \gamma_1^2}{24\sigma_1^2}\right)$;

$$B = \left(1 + \alpha_{OC}(t_x - t)\right)\left(\frac{L_s^2 \gamma_1^2 E_{eq}}{24\sigma_1^2} - \frac{L_s^2 \gamma_1^2}{24\sigma_1} - \sigma_1\right) +$$

$$+ \alpha_{OC} E_{eq}(t_x - t); \; C = \frac{-L_s^2 \gamma_x^2 E_{eq}}{24} .$$

It is possible to solve equation (3) to find the values of the equivalent mechanical stress in the presence and absence of external climatic loads using the Cardano formula.

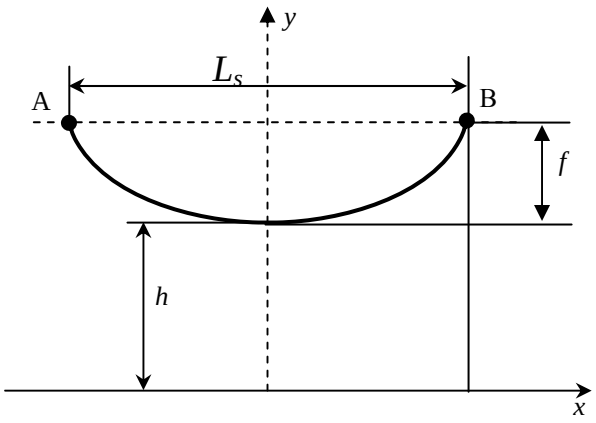



Figure 2 – The sag of the DSOC in the span of the OFOCL with the same height of the suspension points A and B ($L_s$ – the length of the span, $f$ – the arrow of the cable sag, $h$ – the dimensions of the OFOCL in the span)

The magnitude of the sag of the DSOC $f_x$ on the length of the span $L_s$ with the known equivalent mechanical stress under certain external climatic influences $\sigma_x$ and the specific load $\gamma_x$ can be determined by the expression [11]:


DOI 10.15588/1607-3274-2024-4-2

$$f_x = \frac{L_s^2 \gamma_x}{8\sigma_x}. \quad (4)$$

Based on the obtained value of the equivalent mechanical stress $\sigma_x$ in the suspended DSOC, which is under the influence of certain mechanical and physical-climatic loads, it is possible to determine the value of the tensile load of the cable $F_{tlx}$ of the cable according to the expression [11]:

$$F_{tlx} = \sigma_x \cdot S_{SE} \cdot g\ . \quad (5)$$

The last makes it possible to check the correspondence of the $F_{tlx}$ value to the MPTL value of the cable.

According to [5, 11], the permissible equivalent mechanical stress in the DSOC acquires its maximum value under the action of the cable's own weight and the highest or lowest temperature, depending on the type of design of the DSOC. Thus, the construction of the DSOC, which is shown in Fig. 1, has the greatest value of permissible equivalent mechanical stress at the minimum temperature, and the cable design with strength elements made of fiberglass rods – at the maximum.

In the DSOC selected for research, this is due to the presence in the design of elements with different temperature coefficients of linear expansion of materials. Due to the presence of PSE from aramid threads, a tensile stress appears in it at negative temperature, which compensates for the total mechanical compressive stress of all other elements of the cable. In the design of the DSOC with strength elements made of fiberglass rods, the TCLE of the materials of all elements is positive, which explains the change in the equivalent stress of the cable in proportion to the operating temperature.

Since the compatible elements in the design of the DSOC are in close contact with each other, its mechanical properties can be equated to the mechanical properties of a continuous linear element. Then, with relative elongation/compression (deformation), the equivalent mechanical stress of the cable, according to Hooke's law, will be equal to:

$$\sigma_x = \varepsilon_{pcx} \cdot E_{eq}. \quad (6)$$

Thus, with the help of (6), it is possible to calculate $\varepsilon_{pcx}$ and evaluate the deformation processes in the cable and the mechanical state of the OF in it. Therefore, $\varepsilon_{pcx}$, which characterizes the processes of elongation/compression, allows us to estimate the permissible change in the length of the DSOC when acting on it by a longitudinal mechanical force. At the same time, the main task of the OC design is to ensure during operation that no mechanical stress is applied to the optical fibers located in the core of the cable. That is why the DSOC designs with loose tube cable have gained the greatest use in the world. These designs of the cable allow to ensure its permissible relative elongation. It provides additional fiber length in the tube of the optical module before it begins to stretch under the action of tensile forces.

Therefore, the maximum allowable relative elongation of such a cable (MPC) $\varepsilon_{mpc}$ is determined by its design features and the arrangement of OF in it. According to [2], it can be determined by the expression:

$$\varepsilon_{mpc} = \varepsilon_{pc} + \varepsilon_{exOF} + \varepsilon_{pOF}\ . \quad (7)$$

The permissible relative elongation of the cable $\varepsilon_{pc}$ is determined by the expression given, for example, in [2, 5]:

$$\varepsilon_{pc} = -1 + \sqrt{1 + \frac{4 \cdot \pi^2 \cdot R^2}{h_s^2} \cdot \left( \frac{2\Delta R}{R} - \frac{\Delta R^2}{R^2} \right)}. \quad (8)$$

Permissible relative elongation of OF $\varepsilon_{pOF}$ is allowed within the limits of up to 0.25 % for the entire period of operation [2]. As a rule, in practice, during the development of OC $\varepsilon_{exOF}$ is taken equal to 0 in order to obtain a technological reserve for the permissible elongation of the cable.

Additional relative elongation of the cable under certain climatic conditions can be calculated based on expression (6):

$$\varepsilon_{pcx} = \frac{\sigma_x}{E_{eq}}. \quad (9)$$

The actual additional relative elongation of the cable $\varepsilon_{apcx}$, suspended under certain climatic conditions on the OFOCL, can also be determined based on expression (6):

$$\varepsilon_{apcx} = \frac{\sigma_{ax}}{E_{eq}}. \quad (10)$$

The measurement of the actual sag of the DSOC in the span of the OFOCL ($f_{ax}$) during the operation of the line makes it possible to determine the value of its actual mechanical characteristics. Thus, the actual mechanical tension of the cable, based on expression (1), can be determined by the formula:

$$\sigma_{ax} = \frac{L_s^2 \cdot \gamma_1}{8 \cdot f_{ax}}. \quad (11)$$

In turn, the actual tensile load of the cable under certain climatic conditions can be determined by analogy to expression (5) using the formula:

$$F_{atlx} = \sigma_{ax} \cdot S_{SE} \cdot g\ . \quad (12)$$

The value of these mechanical characteristics of the OC will make it possible to estimate the actual elongation


DOI 10.15588/1607-3274-2024-4-2

of the cable and the mechanical state of the optical fiber in it under certain climatic conditions.

Evaluation of the mechanical characteristics of the DSOC and control of the mechanical state of the OF along the length of the span of the OFOCL in the conditions of the given climatic zone require a number of reference data. These data include:

– cable brand and specifications;

– cable design characteristics (cable diameter $D_c$, weight of the cable $P_c$, equivalent temperature coefficient of linear expansion of the cable $\alpha_{OC}$ and Young's modulus of the cable $E_{eq}$, number and type of aramid threads and fiberglass rods from their TCLE, permissible relative elongation of OF in the cable $\varepsilon_{pOF}$ according to the data of the manufacturer's factory);

– the span length of the OFOCL and the value of the sag boom of the DSOC during installation at $t$ °C;

– physical and climatic characteristics of the area of cable operation (wind speed during hoarfrost and ice $\upsilon$, maximum radial thicknesses of hoarfrost and ice walls $\Delta t_h$ and $\Delta t_i$, maximum and minimum temperatures of the climatic area of cable operation);

– corresponding specific loads and mechanical characteristics of DSOC.

In the absence of data from the DSOC developer on $\alpha_{OC}$ and $E_{eq}$, their values are determined according to [12], and the value of $\varepsilon_{pOF}$ is taken as 0.25 %.

Determination of specific loads on DSOC is performed according to [11].

In this work, it is proposed to evaluate the mechanical state of the DSOC during operation by two ways:

– calculating according to expressions (1), (3)–(5) and (9);

– measuring and calculating according to:

a) the results of the measurement of the DSOC sag in the span of the OFOCL ($f_{ax}$) under the influence of $t$ = – 5 ºC, hoarfrost and ice without wind and its calculation ($f_{axc}$) at the temperature $t$ = – 5 ºC, hoarfrost and ice with wind;

b) calculation of its mechanical characteristics ($\varepsilon_{apcx}$, $\sigma_{ac}$ and $F_{atlx}$) according to expressions (10)–(12).

When implementing the methods, the specific load $\gamma_x$ is taken according to reference data of the OFOCL operator.

The last method can be implemented based on the results of the $f_{ax}$ measurement during the maintenance of the OFOCL. Currently, some methods of its measurement using various equipment, including drones and unmanned aerial vehicles, are used.

According to the measured value of $f_{ax}$, there is the actual mechanical stress of the DSOC in the span of the line under the influence of temperature of minus 5 ºC, hoarfrost and ice without wind, and the calculated value of $f_{axc}$ under the influence of temperature of minus 5 ºC and ice with wind. The value of the actual mechanical stress of the cable is found by the expression transformed from expression (4):

$$\sigma_{ax,axc} = \frac{L_s^2 \gamma_x}{8 f_{ax,axc}}. \qquad (13)$$

Determination of the permissible relative elongation of the cable is carried out according to expression (10), and its tensile load – by (5).

Thus, the results of the evaluation of the mechanical characteristics of the DSOC during operation, obtained by the first way, should be considered as reference for the technical personnel of the OFOCL operator.

The results of the assessment of the mechanical characteristics of the DSOC obtained by the second way should be considered as actual, that is, those that must be used to monitor the mechanical state of the optical fiber.

This way was developed on the basis of an assessment of the actual mechanical characteristics of the DSOC during operation and the results of experimental studies:

– measurement of the actual cable sag in the span of the OFOCL and calculation of its value when the cable is affected by the temperature of $t$ = – 5 ºC, the weight of ice with wind pressure;

– calculation of the actual mechanical stress, additional elongation and tensile load of the cable at different span lengths.

The last made it possible to determine and establish the presence/absence of excessive tensile forces on the optical fiber in the cable.

According to the results of the calculation of the maximum permissible elongation of the cable (7), its permissible relative elongation – expression (8) – with the values of the permissible elongation of the OF and its excess length in the OM tube, according to the data of the developer of DSOC, it is possible to assess the mechanical state of the OF in the researched cable of the brand OKL-3-D2A14. Therefore, at the determined value of $\varepsilon_{mpc}$, a limit is obtained, i.e., the maximum elongation of the DSOC, the value of which, when compared with the elongation of the cable under certain mechanical and physical-climatic conditions, makes it possible to control the mechanical state of the optical fiber.

Therefore, the values of $\varepsilon_{mpc}$ and $\varepsilon_{apcx}$ obtained during the operation of DSOC create the possibility of control the mechanical state of both cable and optical fibers. The last makes it possible to monitor them and take preventive measures at the OFOCL to ensure its service life.

The value of the cable sag in the span of the OFOCL under the influence of the temperature of minus 5 ºC, ice with wind ($f_{7i}$) is calculated according to the parabolic curve of the dependence of $f_{ax}$, $f_{axc}$ on the specific loads ($\gamma_{3h}$, $\gamma_{3i}$ and $\gamma_{7h}$) of the OFOCL. An example of calculation $f_{7i}$ is given in next chapter. Studies of the theoretical foundations of the method of control the mechanical state of OF were carried out with the help of software implementation in the Python environment in accordance with the graphic model presented in Fig. 3.


DOI 10.15588/1607-3274-2024-4-2

## 4 EXPERIMENTS

Control of the mechanical state of OF and DSOC under the influence of current mechanical, physical and climatic conditions of operation requires:

1. Calculation of specific loads to DSOC.
2. Calculation of the equivalent mechanical stress of DSOC.

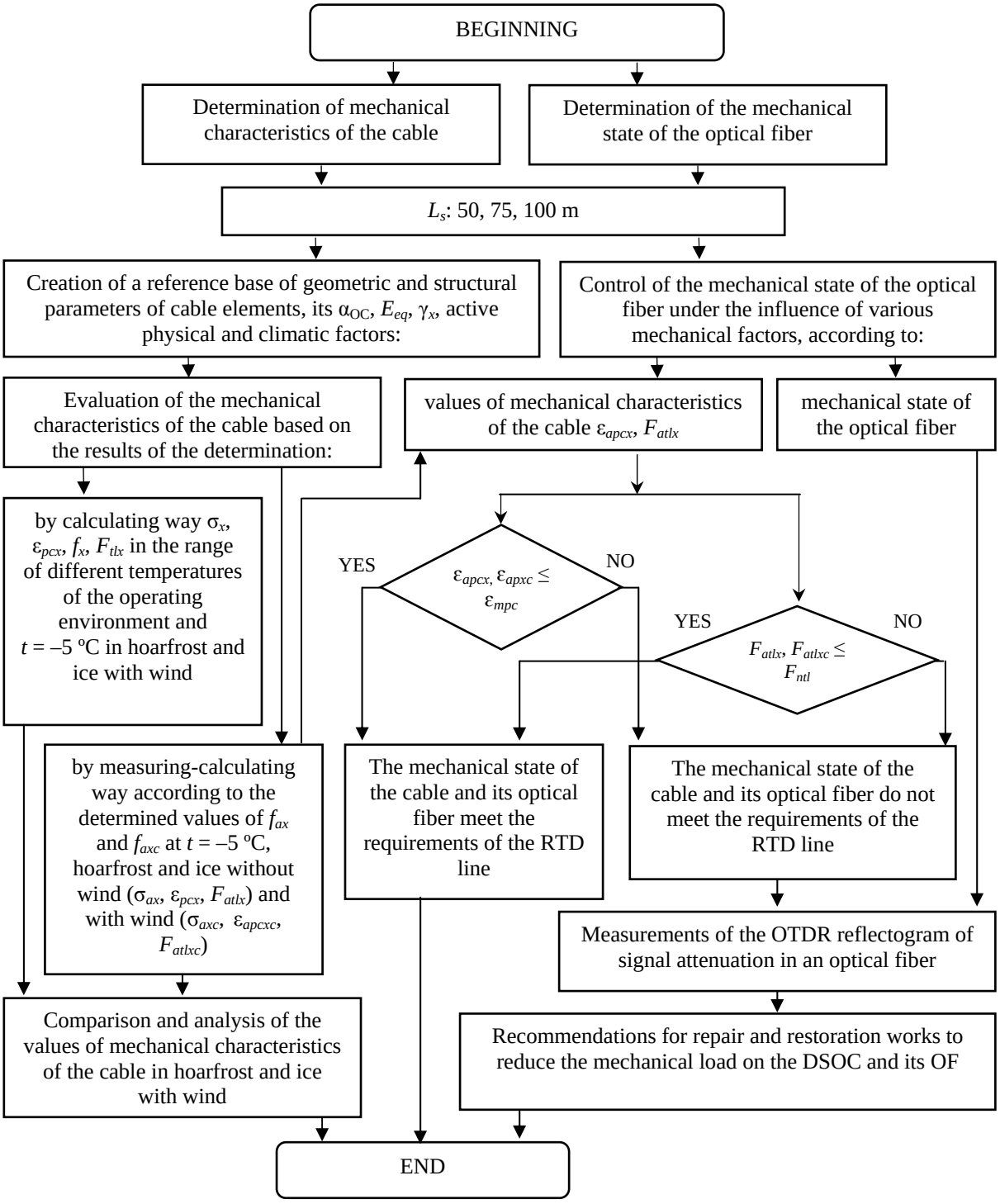



Figure 3 – Graphic model of the study by the method of control the mechanical state of the optical fiber depending on the cable brand and the climatic zone of operation


DOI 10.15588/1607-3274-2024-4-2

3. Determination of the arrow sag of the DSOC in the span. Comparison of the obtained value $f_x$ with the measured actual cable sag. In the case of a difference in results, make an appropriate analysis of the reasons for the change in cable sag from the expected value and foresee the possibility of implementing preventive measures.

4. Determination of the tensile load acting on the cable and its comparison with the value of MPTL of DSOC, which is specified in its technical characteristics (or technical conditions for the cable). In case of exaggeration (65...70) % of the MPTL, make an appropriate analysis of the reasons for cable overload in order to ensure its mechanical integrity and foresee the possibility of implementing preventive measures.

5. Determination of the actual effective relative elongation of the cable ($\varepsilon_{apc}$) and its value according to expression (11) to estimate the mechanical state of the fiber.

As an example, in this paper, the specific loads from mechanical and physical-climatic influences and the corresponding mechanical characteristics of the DSOC type OKL-3-D2A14 made by PJSC "Odeskabel" were determined based on the ways of evaluating the mechanical characteristics of the DSOC during operation. In the calculations it was assumed: the diameter of the cable $D_c$ = = 13.6 mm, the specific weight of the cable $P_c$ = = 155 kg/km, the equivalent TCLE of the cable $\alpha_{OC}$ = = $6.226 \cdot 10^{-6}$ $K^{-1}$, the equivalent Young's modulus of the entire cable ($E_{eq}$ = 20010 N/mm$^2$, determined according to [12]), the conditions of the Odesa climate zone (Black Sea region) – wind speed during hoarfrost and ice $V$ = = 28.3 m/s, the maximum radial thickness of the hoarfrost wall on the cable $\Delta t_h$ = 0.5 mm and the ice wall $\Delta t_i$ = = 28 mm, the initial sag of the DSOC at the temperature of 20 ºC is equal to 1 % of the span length, the span length of the OFOCL 50 m, 75 m and 100 m. As PSE in the calculations aramid threads of the type "Twaron D1052 8050" with TCLE $\alpha_{at}$ = – $3 \cdot 10^{-6}$ 1º $K^{-1}$ and a fiberglass rod of the type "Polystal P20" with TCLE $\alpha_{fr}$ = =$6.6 \cdot 10^{-6}$ 1º $K^{-1}$ are used.

Determination of the reference data of OFOCL was carried out in accordance with the theoretical foundations of the method and ways of assessing the mechanical characteristics of DSOC in the Odesa climatic zone (chapter 3).

In Odesa climate zone, the air temperature varies from maximum of plus 37 °C to minimum of minus 28 °C. Due to the fact that the cable has a black sheath, the maximum temperature was assumed to be plus 50 °C in the calculations of its specific load and mechanical characteristics.

Let's perform the evaluation of mechanical characteristics of DSOC by calculating method.

The results of the calculation of the specific loads on the cable and its mechanical characteristics during operation in the Odesa climatic zone according to this way of assessment for different span lengths and the effect of mechanical and various physical and climatic factors are given in the Table 1.

The results of calculating the values of the mechanical characteristics of DSOC, which are given in Table 1, shown that they increase, primarily, from a change in temperature from plus 50 °C to minus 28 °C. Their worst values appear at $t$ = –5 °C during hoarfrost and ice. In turn, changes in the values of the mechanical characteristics of DSOC brand OKL-3-D2A14 in absolute value (Table 1) are given in Table 2. These characteristics were calculated for three values of the length of the line span: 50 m, 75 m and 100 m.

Let's perform the evaluation of mechanical characteristics of DSOC by measuring-calculating way.

The results of determining the mechanical characteristics of DSOC during operation in the selected climatic zone by the measuring-calculating way are given in Table 3. At the same time, the values of the arrow sag $f_x$ were taken based on the results of two measurements at the temperature $t$ = –5 °C and the effect of hoarfrost and ice without wind on span lengths 50 m, 75 m and 100 m. The value of the cable arrow sag at the temperature $t$ = –5 ° C, hoarfrost and ice with wind were calculated according to the parabolic curve of dependence of $f_{ax}$, $f_{axc}$ on the specific loads of the cable: $\gamma_{3h}$ – hoarfrost, $\gamma_{3i}$ – ice and $\gamma_{7h}$ – hoarfrost. The value of $f_x$ at $\gamma_{7h}$ was found according to the data in the Table 1. The dependence of the cable sag arrow of the brand OKL-3-D2A14 at $L_s$ = 100 m on the specific loads of the OFOCL are shown in Fig. 4.

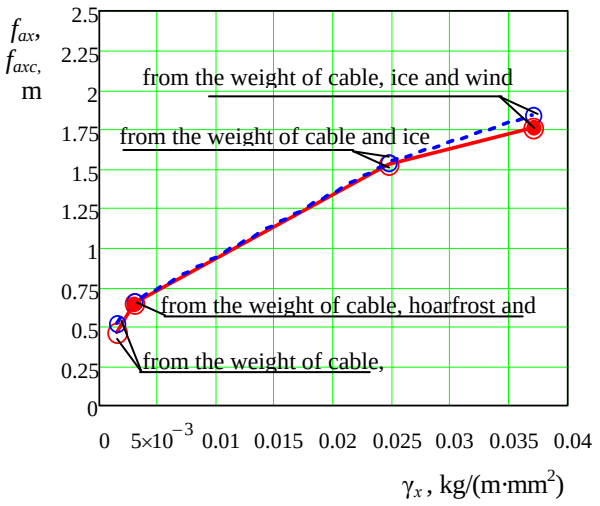



Figure 4 – Dependence of the cable arrow sag at $L_s$ = 100 m on the specific loads of OFOCL, obtained by —— calculating and – – measuring-calculating ways ( • – $f_{ax}$ , ○ – $f_{axc}$)

Table 3 shows the results of determining the mechanical characteristics of the DSOC ($\sigma_{ax}$, $\varepsilon_{apcx}$, $F_{atlx}$), measurements and calculations of the cable sag in spans with different lengths ($f_{ax}$, $f_{axc}$) during its operation in the selected climatic zone when acting on the OFOCL hoarfrost and ice without wind.

As noted earlier, this way of assessing the mechanical characteristics of DSOC gives their actual values under the conditions of the location of the line under the influence of hoarfrost, ice without wind. Therefore, in order to


DOI 10.15588/1607-3274-2024-4-2

compare the results of determining the mechanical characteristics of the cable according to the first and second ways, an analysis of the data in the Table 1 and Table 3 was carried out. This analysis shown that the difference in the values of mechanical characteristics obtained by these ways is from 3.98 % to 4.50 %.

Let's perform the evaluation of the method of control the mechanical state of the optical fiber.

As noted earlier, certain physical and climatic loads on the DSOC cause its longitudinal deformations, which lead to the elongation of only the cable structure and/or its elongation together with the optical fiber. Excessive elongation of DSOC can lead to decreasing the size of the line and exceeding the norms of the cable parameters according to the Technical Specifications. The last, with a significant and long-term effect of influential factors, leads to a difference between $F_{atlx}$ and $F_{mtl}$, which can cause a breakage of the OF and a complete failure of the cable. Thus, the work establishes a clear mathematical toolkit, which allows to control the mechanical state of the optical fiber and the mechanical characteristics of the cable as a whole under the influence of physical and climatic influences of the operating environment.

As an example, in the study of the method of control the mechanical state of an optical fiber, the values of the actual effective elongation of the fiber and the actual tensile load on the cable were determined using expressions (10) – (12). The values of $\varepsilon_{apcx}$ and $F_{atlx}$ were obtained when the temperature changed from plus 20 °C to minus 5 °C, hoarfrost and ice according to chapter 3. When studying these mechanical characteristics at $t$ = +20 °C, the cable's own weight was taken into account, and at $t$ = = –5 °C its own weight, the weight of hoarfrost and ice, and the wind pressure. For the analysis of the mechanical state of OF under the above conditions, $\varepsilon_{mpcx}$ was calculated according to the reference data. At the same time, using expression (8), it was determined that for the studied design of the DSOC, the permissible relative elongation of the cable is 0.34 %. Then $\varepsilon_{mpc}$ according to (7) with $\varepsilon_{exOF}$ = 0 and $\varepsilon_{pOF}$ = 0.25 % is 0.59 %.

Table 1 – The value of the specific loads $\gamma_x$ on the cable brand OKL-3-D2A14 and its mechanical characteristics, obtained by the calculating way

| № | $\gamma_x$, kg/(m·mm$^2$), in the Odesa climate zone under the influence of various factors | $L_{sx}$, m | Mechanical characteristics of DSOC along the length of the line span | | | |
|---|---|---|---|---|---|---|
| | | | $\sigma_x$, kg/mm$^2$ | $\varepsilon_{pcx}$, % | $f_x$, m | $F_{tlx}$, N |
| 1 | weight of cable, $t$ = +50 °C, $\gamma_1$ = 0.00130 | 50 | 0.675 | 0.033 | 0.602 | 181.8 |
| | | 75 | 1.040 | 0.051 | 0.879 | 280.1 |
| | | 100 | 1.416 | 0.069 | 1.147 | 381.4 |
| 2 | weight of cable, $t$ = +20 °C, $\gamma_1$ = 0.00130 | 50 | 0.813 | 0.040 | 0.500 | 218.8 |
| | | 75 | 1.219 | 0.060 | 0.750 | 328.2 |
| | | 100 | 1.625 | 0.080 | 1.000 | 437.6 |
| 3 | weight of cable, $t$ = –28 °C, $\gamma_1$ = 0.00130 | 50 | 1.150 | 0.056 | 0.353 | 309.6 |
| | | 75 | 1.600 | 0.078 | 0.571 | 430.9 |
| | | 100 | 2.037 | 0.100 | 0.798 | 548.5 |
| 4 | weight of cable and hoarfrost, $t$ = –5 °C, $\gamma_{3h}$ = 0.00145 | 50 | 1.018 | 0.050 | 0.445 | 274.0 |
| | | 75 | 1.462 | 0.072 | 0.697 | 393.8 |
| | | 100 | 1.896 | 0.093 | 0.956 | 510.5 |
| 5 | wind pressure at hoarfrost, $t$ = –5 °C, $\gamma_{5h}$ = 0.00147 | 50 | 1.024 | 0.050 | 0.449 | 275.7 |
| | | 75 | 1.470 | 0.072 | 0.703 | 395.9 |
| | | 100 | 1.905 | 0.093 | 0.965 | 512.9 |
| 6 | weight of cable, hoarfrost and wind pressure, $t$ = –5 °C, $\gamma_{7h}$ = 0.00290 | 50 | 1.444 | 0.071 | 0.628 | 388.7 |
| | | 75 | 1.999 | 0.098 | 1.020 | 538.2 |
| | | 100 | 2.522 | 0.124 | 1.437 | 679.2 |
| 7 | weight of cable and ice, $t$ = –5 °C, $\gamma_{3i}$ = 0.02464 | 50 | 5.260 | 0.258 | 1.465 | 1415.0 |
| | | 75 | 6.970 | 0.342 | 2.486 | 1877.0 |
| | | 100 | 8.520 | 0.417 | 3.617 | 2293.0 |
| 8 | weight of cable, ice and wind pressure, $t$ = –5 °C, $\gamma_{7i}$ = 0.03700 | 50 | 6.827 | 0.335 | 1.694 | 1838.0 |
| | | 75 | 9.026 | 0.442 | 2.882 | 2430.0 |
| | | 100 | 11.005 | 0.540 | 4.203 | 2963.0 |

Table 2 – Changes of the values $\sigma_x$. $\varepsilon_{pcx}$. $f_x$ and $F_{tlx}$ DSOC in absolute value at the temperature $t$ = +50 °C relative to their values at $t$ = –5 °C in hoarfrost and ice with wind obtained by the calculating way

| № | $L_{sx}$, m | Changing the values of the mechanical characteristics of the cable | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | characteristic name | | | | | | | |
| | | $\sigma_x$, kg/mm$^2$, at: | | $\varepsilon_{pcx}$, %. at: | | $f_x$, m, at: | | $F_{tlx}$, kN, at: | |
| | | hoarfrost | ice | hoarfrost | ice | hoarfrost | ice | hoarfrost | ice |
| 1 | 50 | 0.769 | 6.152 | 0.038 | 0.302 | 0.026 | 1.092 | 0.207 | 1.656 |
| 2 | 75 | 0.959 | 7.986 | 0.047 | 0.391 | 0.141 | 2.003 | 0.258 | 2.150 |
| 3 | 100 | 1.106 | 9.589 | 0.055 | 0.471 | 0.290 | 3.056 | 0.298 | 2.581 |


DOI 10.15588/1607-3274-2024-4-2

Table 3 – Values of the mechanical characteristics of OKL-3-D2A14 DSOC, obtained by the measuring-calculating way under the influence of the weight of the cable, hoarfrost, ice without wind pressure and temperature $t$ = –5 °C

| № | $L_s$, m | $f_{ax}$, m, at: | | Values of the mechanical characteristics of the cable | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $\sigma_{ax}$, kg/mm$^2$, at: | | $\varepsilon_{ax}$, %, at: | | $F_{atlx}$, N, at: | |
| | | hoarfrost | ice | hoarfrost | ice | hoarfrost | ice | hoarfrost | ice |
| 1 | 50 | 0.463 | 1.524 | 0.979 | 5.052 | 0.048 | 0.248 | 263.5 | 1361 |
| 2 | 75 | 0.725 | 2.585 | 1.406 | 6.702 | 0.069 | 0.329 | 378.7 | 1805 |
| 3 | 100 | 0.994 | 3.762 | 1.823 | 8.187 | 0.089 | 0.401 | 491.0 | 2205 |

In addition, an analysis of the mechanical state of the optical fiber is provided (Table 4) for line span lengths of 50, 75, and 100 m.

The results of the analysis of the mechanical state of the optical fiber in the investigated structure of the DSOC under the influence of physical and climatic influences of the operating environment are given in Table 4, Fig. 5 and Fig. 6.

Data analysis (Table 4) shown that when:

– the maximum permissible relative elongation of the DSOC $\varepsilon_{mpc}$ = 0.59 % on all lengths of the span, the additional relative elongation of the cable is less than this value;

– the length of the span of the overhead line 100 m, the actual tensile load of the cable $F_{atlxc}$ = 2.722 kN is greater than the norm of its permissible tensile load according to [7] $F_{ntl}$ = 2.600 kN.

Therefore, the cable brand OKL-3-D2A14 in the Odesa climate zone at $L_s$ = 100 m cannot be used on the OFOCL.

Table 4 – Results of the analysis of the mechanical state of the optical fiber in DSOK OKL-3-D2A14 under the influence of physical and climatic influences at different lengths of span at the temperature of plus 20 °C and minus 5 °C

| № | $\gamma_x$, kg/(m·mm$^2$), in the Odesa climate zone under the influence of various factors | Mechanical characteristics of DSOC on the length of the span | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\varepsilon_{apc}$, $\varepsilon_{apcxc}$, %, at span lengths of the OFOCL, m | | | | | | $F_{atl}$, $F_{atlxc}$, N, at span lengths of the OFOCL, m | | | | | | Mechanical state of the OF at the span length of the OFOCL, m | | |
| | | 50 | | 75 | | 100 | | 50 | | 75 | | 100 | | 50 | 75 | 100 |
| | | $\varepsilon_{apc}$, $\varepsilon_{apcxc}$ | $\varepsilon_{pc}$ | $\varepsilon_{apc}$, $\varepsilon_{apcxc}$ | $\varepsilon_{pc}$ | $\varepsilon_{apc}$, $\varepsilon_{apcxc}$ | $\varepsilon_{pc}$ | $F_{atl}$, $F_{atlxc}$ | $F_{ntl}$ | $F_{atl}$, $F_{atlxc}$ | $F_{ntl}$ | $F_{atl}$, $F_{atlxc}$ | $F_{ntl}$ | | | |
| 1 | $t$ = 20 °C and cable weight, $\gamma_1$ = 0,00130 | 0.040 < 0.340 | | 0.060 < 0.340 | | 0.080 < 0.340 | | 218.80< 2600.0 | | 328.20 < 2600.0 | | 437.60 < 2600.0 | | OF straightens in the core and is not subject to longitudinal elongation | | |
| 2 | $t$ = –5 °C, weight of cable and hoarfrost, $\gamma_{3h}$ = 0,00145 | 0.048 < 0.340 | | 0.069 < 0.340 | | 0.089 < 0.340 | | 263.50 < 2600.0 | | 378.70 < 2600.0 | | 491.00 < 2600.0 | | OF straightens in the core and is not subject to longitudinal elongation | | |
| 3 | $t$ = –5 °C, wind pressure at hoarfrost, $\gamma_{5h}$ = 0,00147 | 0.050 < 0.340 | | 0.072 < 0.340 | | 0.093 < 0.340 | | 275.70 < 2600.0 | | 395.90 < 2600.0 | | 512.90 < 2600.0 | | OF straightens in the core and is not subject to longitudinal elongation | | |
| 4 | $t$ = –5 °C, cable weight, hoarfrost, wind, $\gamma_{7h}$ = 0,00290 | 0.068 < 0.340 | | 0.097 < 0.340 | | 0.125 < 0.340 | | 370.90 < 2600.0 | | 531.50 < 2600.0 | | 685.00 < 2600.0 | | OF straightens in the core and is not subject to longitudinal elongation | | |
| 5 | $t$ = –5 °C, weight of cable and ice, $\gamma_{3i}$ = 0,02464 | 0.248 < 0.340 | | 0.329 < 0.340 | | 0.401 > 0.340 | | 1361.0 < 2600.0 | | 1805.0 < 2600.0 | | 2205.0 < 2600.0 | | OF straightens in the core and is not subject to longitudinal elongation | | OF is subject to elongation by 0.06 % |
| 6 | $t$ = –5 °C, weight of cable and ice, wind pressure, $\gamma_{7i}$ = 0,03700 | 0.31 < 0.340 | | 0.418 > 0.340 | | 0.496 > 0.340 | | 1702.0 < 2600.0 | | 2298.0 < 2600.0 | | 2722.0 > 2600.0 | | OF straighstens in the core | OF is subject to elongation by 0.08 % | OF is subject to elongation by 0.16 % |


DOI 10.15588/1607-3274-2024-4-2

## 5 RESULTS

In the work, the method of control the mechanical state of the optical fiber of DSOC during operation was created. Based on the developed theoretical basis of the method, experimental studies were performed to evaluate the mechanical characteristics of the cable OKL-3-D2A14 at three span lengths of the line located in the Odesa climatic zone and the mechanical state of its fiber. On the basis of the developed reference data of the method and measurements and calculations of the cable arrow sag, an analysis of the results of the study of ways for evaluating the mechanical characteristics of cables was carried out. These results proved the possibility of using the measuring-calculating way at controlling the mechanical state of the optical fiber during the operation of the OFOCL.

The results of a comparison of the measured values of the sag of the DSOC (Table 3) and those calculated under the influence of the weight of the cable, the weight of hoarfrost and ice without wind pressure (Table 1) shown that the measured values are greater than the calculated values by up to 4 % both in hoarfrost and in ice at all span lengths.

The analysis of the obtained values of mechanical characteristics of DSOC by these methods shown that their difference is also up to 4 %.

The obtained values of the actual sag of the cable in the worst operating conditions during icing with wind (Table 4) allow us to determine the dimensions of the line and the value of the tensile load of the DSOC.

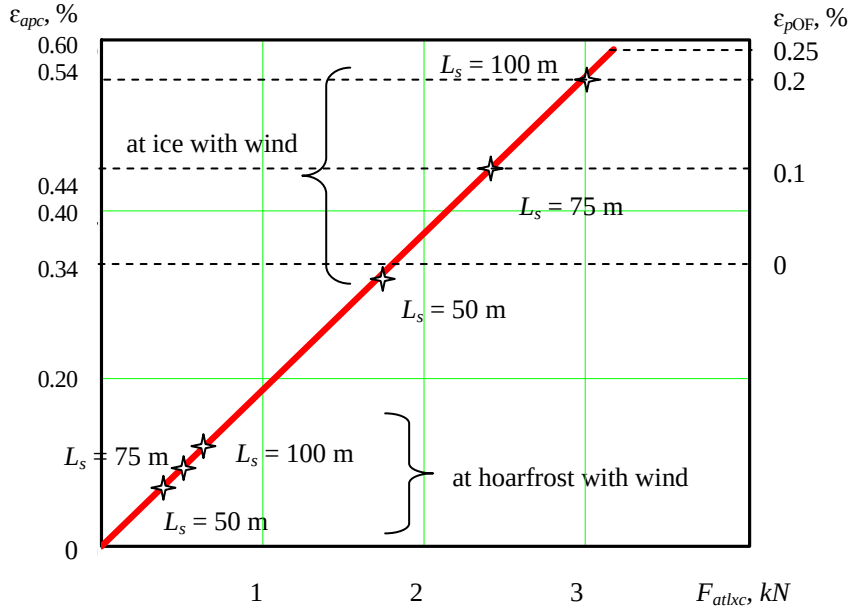



Figure 5 – Assessment of the mechanical state of OF under the influence of physical and climatic influences

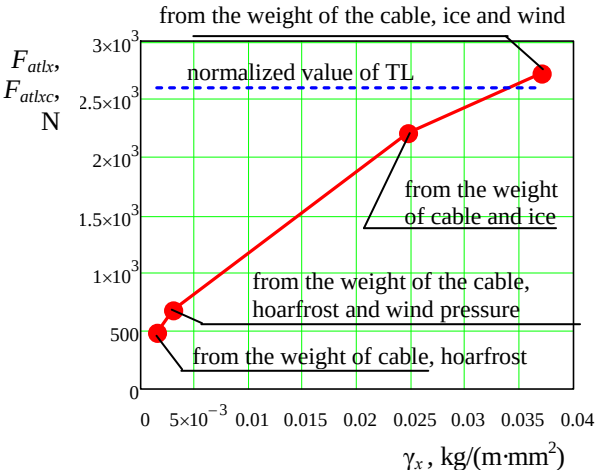



Figure 6 – Dependence of the tensile load of the cable on the specific loads of the OFOCL on the length of the span of 100 m

The developed method of control the mechanical state of OF shown that:

a) for all span lengths and temperatures from plus 20 °C to minus 5 °C during hoarfrost with wind, the optical fiber is not subject to longitudinal elongation;

b) at temperature $t$ = –5 °C:

– and at ice without wind, the optical fiber is straightened in the tube of the optical module at $L_s$ = 50 m, 75 m, and at $L_s$ = 100 m it is subject to elongation by 0.06 %;

– and ice with wind, the optical fiber is straightened in the tube of the optical module at $L_s$ = 50 m, and at $L_s$ = 75 m it is subject to elongation by 0.08 % and at $L_s$ = 100 m – by 0.16 %.

So, the research results proved that the cable brand OKL-3-D2A14 in the Odesa climate zone on lines with a span length of 100 m or more is inadmissible for operation.

## 6 DISCUSSION

The approaches presented in the work in the created method are confirmed by their necessity, given in the international and Ukrainian regulatory and technical documentation for the building and operation of the OFOCL.

The method is cheap and easily accessible to operators of overhead communication lines. Therefore, its use improves the monitoring of the mechanical state of the pipeline and leads to a reduction in the costs of its operation.


DOI 10.15588/1607-3274-2024-4-2

The proposed method of control the mechanical state of the optical fiber in DSOC allows:

– choose the brand of cable when designing the OFOCL for hanging on various engineering structures in the given climatic zone and the length of the span with the same height of the cable suspension points;

– monitor the technical state of the line based on the results of the assessment of the mechanical characteristics of the cable and control of the mechanical state of the optical fiber and measurements of its attenuation coefficient;

– the studies carried out in the work allow to ensure the specified service life of the lines with proper maintenance or repair and restoration works of DSOC. This is confirmed by the results of the experiment performed in the work. Therefore, the method will help the service personnel of the lines to take timely measures to eliminate problems in the mechanical states of the cable and its fiber.

## CONCLUSIONS

The research carried out in this work made it possible to draw the following conclusions:

1. The international and domestic regulatory and technical documentations of the OFOCL and a number of scientific works regulate the control of the mechanical characteristics of the optical cable during operation, but the control of the technical state of the fiber in full to ensure the quality and reliability of the line during the specified service life is not provided.

2. In this work, the method of control the mechanical state of the optical fiber of dielectric self-supporting optical cables during operation has been developed due to the developed ways of evaluating the mechanical characteristics of the cable along the length of the span under the influence of mechanical, physical and climatic factors of the area where the line is located.

3. With the help of the developed method, the determination of mechanical stresses and the study of additional relative elongations of the cable and its tensile loads in the span of the line during operation on the OFOCL under the conditions of various mechanical and physical-climatic loads (ice, wind pressure, the affect of temperature and their combination) of the Odesa climatic zones at three span length values. It was established that changes in mechanical and physical-climatic loads of operating conditions cause different magnitudes of longitudinal deformations of DSOC and OF as a whole. Using this method, it was established that the cable brand OKL-3-D2A14 with the specified design parameters at a span length of 100 m under the worst climatic conditions ensures the mechanical integrity of the OF with an additional relative elongation of 0.16 %, but in terms of actual tensile strength, it does not meet the norm – 2.6 kN.

As the mechanical state of cable and optical fiber deteriorates over time, it is very important to monitor them to prevent fiber breakage.

4. The proposed method of control the mechanical state of the optical fiber on the suspended DSOC on the OFOCL allows monitoring changes in the mechanical state of optical fibers and the cable as a whole to prevent the additional appearance of unacceptable tensile forces and failure of the communication line. This method can be used for application by the relevant departments for the technical operation of telecommunication lines on communication networks based on suspended optical cables.

## ACKNOWLEDGEMENTS

The work was carried out at the State University of Intelligent Technologies and Telecommunications (former O.S. Popov Odesa National Academy of Telecommunications) during 2015 – 2022 y. as part of a number of initiative research works: “Methods and devices for increasing the efficiency of transmission lines. Stage 3 – 5” (state registration number 0117U003419) and “Methods of increasing the reliability and bandwidth of fiber-optic communication lines” (state registration number 0119U103658).

## REFERENCES

1. IEEE Standard for Testing and Performance for All-Dielectric Self-Supporting (ADSS) Fiber Optic Cable for Use on Electric Utility Power Lines, *IEEE Standard 1222*, 2019. DOI: 10.1109/IEEESTD.2020.9052820.
2. Günter B. The global cable industry: materials, markets, products. Sidney, John Wiley & Sons, 2021. – 416 p. DOI: 10.1002/9783527822263.
3. Bondarenko O. V., Stepanov D. M., Verbytskyi O. O., Siden S. V. Method of evaluation the efficiency of fiber-optic cables models with multi-modular design based on mass and dimensional indicators, *Radio Electronics, Computer Science, Control*, 2024, № 1, pp. 6–16. DOI: 10.15588/1607-3274-2024-1-1.
4. Bailey D., Wright E. Fiber optic cable construction [Electronic resource]. Access mode: https://www.sciencedirect.com/book/9780750658003/practical-fiber-optics#book-info. DOI: https://doi.org/10.1016/B978-0-7506-5800-3.X5000-0.
5. Gunes I. Examination of wind effect on ADSS cables aging test, *Electrica*, 2018, Vol. 18, No. 2. pp. 321–324. DOI: 10.5152/iujeee.2018.1818.
6. Tingting Z., Minning J. Optimization of the optical fiber with triple sector stress elements, *Optical Fiber Technology*, 2020, Vol. 57. Article no. 102212. DOI: 10.1016/j.yofte.2020.102212.
7. IEEE Guide for Overhead Alternating Current (AC) Transmission Line Design, *IEEE Standard 1863*, 2024. DOI: 10.1109/IEEESTD.2024.10622047.
8. Rekomendaciï z pidvishuvannja optichnih kabeliv na oporah povitrjanih linij zv’jazku, LEP, kontaktnoï merezhi zaliznic’: R 45-010-2002. [Chinnij vid 27.04.2003 r.]. K., Derzhavnij komitet zv’jazku ta informatizaciï Ukraïni, 2004, 95 p.
9. Canudo J., Sevillano P., Iranzo A., Kwik S., Preciado-Garbayo J. and J. Subías. Overhead Transmission Line Sag Monitoring Using a Chirped-Pulse Phase-Sensitive OTDR, *IEEE Sensors Journal*, 2024, January, Vol. 24, No. 2, pp. 1988–1995. DOI: 10.1109/JSEN.2023.3340296.
10. Mekhtiyev A., Bulatbayev F., Neshina Y., Siemens E. The external mechanical effects on the value of additional losses in the telecommunications fiber optic cables under operating conditions, *Elektrotechnik, Maschinenbau und Wirtschaftsingenieurwesen*, 2018. Vol. 6(1). pp. 123–127. DOI:10.13142/kt10006.43.
11. Linijno-kabel’ni sporudi telekomunikacij. Proektuvannja : GBN B.2.2-34620942-002:2015. [Chinnij vid 01.08.2015 r.]. K., Administracija Derzhspeczv’jazku, 2015, 140 p.
12. Feng X., Changsen S., Xiaotan Z., Farhad A. Determination of the coefficient of thermal expansion with embedded long-gauge fiber optic sensors, *Measurement Science and Technology*, 2010, Vol. 21, No.6. Article no. 065302. DOI:10.1088/0957-0233/21/6/065302.

Received 01.08.2024.
Accepted 17.10.2024.


DOI 10.15588/1607-3274-2024-4-2

УДК 681.7.068

# МЕТОД КОНТРОЛЮ МЕХАНІЧНОГО СТАНУ ОПТИЧНОГО ВОЛОКНА ДІЕЛЕКТРИЧНОГО САМОУТРИМНОГО ОПТИЧНОГО КАБЕЛЮ ПРИ ЕКСПЛУАТАЦІЇ

**Бондаренко О. В.** – д-р техн. наук, професор, Лауреат Державної Премії України в галузі науки і техніки, Одеса, Україна.
**Степанов Д. М.** – канд. техн. наук, доцент, доцент кафедри комутаційних систем електронних комунікацій, Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна.

## АНОТАЦІЯ

**Актуальність.** Нормативно-технічна документація для волоконно-оптичних ліній зв'язку (ВОЛЗ) регламентує контроль механічного стану оптичного кабелю (ОК) під час експлуатації, але контроль механічного стану волокна в повній мірі для забезпечення якості та надійності лінії протягом заданого терміну служби не передбачає.

Як відомо, що для забезпечення надійності оптичного кабелю, як правило, допустиме видовження оптичного волокна (ОВ) складає $\varepsilon_{\text{доВ}} < (0{,}2…0{,}25)$ %, прийняте при конструюванні кабелю. Але при експлуатації в волокнах можлива поява багаторазових надлишкових видовжень, що перевищують ці значення.

Таким чином, розробка та обґрунтування способів оцінки механічних характеристик діелектричного самоутримного оптичного кабелю (ДСОК) та методу повного контролю механічного стану оптичного волокна являється необхідним. Останнє може привести до дострокового виходу із ладу оптичного волокна.

**Мета.** Розробка та обґрунтування методу контролю механічного стану оптичних волокон підвісних ДСОК, а також оцінка умов деформації оптичних волокон в їх осерді з появою поздовжніх розтягувальних/стискальних навантажень при експлуатації.

**Метод.** Розроблено та запропоновано два способи оцінки механічних характеристик ДСОК та метод контролю механічного стану його волокна. Для цього в роботі прийняті такі характеристики кабелю та волокна: відносне видовження кабелю та волокна ($\varepsilon_{\text{кх}}$, $\varepsilon_{\text{ОВх}}$), довжина прольоту ($L_{\text{прх}}$) лінії, стріла провисання кабелю в прольоті ($f_x$) та розтягувальне навантаження (РН) кабелю ($F_{\text{рнх}}$), що обумовлює поздовжню деформацію ε. При цьому в методі запропоновано контроль механічного стану оптичного волокна при експлуатації ДСОК виконувати за рахунок визначення його діючого відносного видовження за механічними та фізико-кліматичними умовами розташування лінії.

В роботі запропоновано, завдяки розробленим довідковим даним для вибору типу кабелю та кліматичної зони розташування лінії, вимірювальному обладнанню та математичному інструментарію, вимірювати та розраховувати такі механічні характеристики:

– еквівалентну механічну напругу в ДСОК;
– розрахункову та фактичну стріли провисання кабелю в прольоті;
– фактично діюче відносне видовження кабелю;
– фактичне розтягувальне навантаження, що діє на кабель.

Приведені в роботі способи та метод дозволяють при експлуатації ДСОК здійснювати повну оцінку механічних характеристик кабелю та контроль механічного стану оптичного волокна. Це створює можливість моніторингу їх змін для недопущення появи надмірних навантажень при експлуатації та виходу з ладу волоконно-оптичної лінії зв'язку.

Даний метод можливо рекомендувати для застосування відповідними підрозділами з технічної експлуатації телекомунікаційних ліній та мереж зв'язку на базі підвісних оптичних кабелів.

**Результати.** В роботі приведені результати розробки та обґрунтування методу контролю механічного стану оптичних волокон діелектричних самоутримних оптичних кабелів при експлуатації. Для прикладу, використовуючи розроблений метод, показано, що в кабелі марки ОКЛ-3-Д2А14 виробництва ПАТ «Одескабель» в умовах Одеської кліматичної зони (Причорноморського регіону) оптичні волокна при довжині прольоту лінії 100 м підлягають видовженню на 0,16 %, а ДСОК підлягає фактичному розтягувальному зусиллю – 2,722 кН. Цей результат контролю механічного стану ОВ встановив, що такий прольот лінії забезпечує його механічну цілісність у межах допустимої деформації 0,25 %, прийнятої при конструюванні кабелю, але перевищує його допустиме розтягувальне навантаження 2,6 кН.

**Висновки.** Наукова новизна результатів роботи полягає в тому, що вперше розроблено способи повної оцінки механічних характеристик ДСОК при експлуатації та метод повного контролю механічного стану його оптичного волокна. Він дозволяє виконувати моніторинг змін механічного стану оптичного волокна кабелю.

**КЛЮЧОВІ СЛОВА:** відносне видовження, діелектричний самоутримний оптичний кабель, оптичні волокна, поздовжнє розтягувальне навантаження, фізико-кліматичні навантаження.

## ЛІТЕРАТУРА

1. IEEE Standard for Testing and Performance for All-Dielectric Self-Supporting (ADSS) Fiber Optic Cable for Use on Electric Utility Power Lines. – IEEE Standard 1222, 2019. DOI: 10.1109/IEEESTD.2020.9052820.
2. Günter B. The global cable industry: materials, markets, products / B. Günter. – Sidney : John Wiley & Sons, 2021. – 416 p. DOI: 10.1002/9783527822263.
3. Method of evaluation the efficiency of fiber-optic cables models with multi-modular design based on mass and dimensional indicators / [O. V. Bondarenko, D. M. Stepanov, O. O. Verbytskyi, S. V. Siden] // Radio Electronics, Computer Science, Control. – 2024. – No. 1. – P. 6–16. DOI: 10.15588/1607-3274-2024-1-1.
4. Bailey D. Fiber optic cable construction [Electronic resource] / D. Bailey, E. Wright. – Access mode: https://www.sciencedirect.com/book/9780750658003/practical-fiber-optics#book-info. DOI: https://doi.org/10.1016/B978-0-7506-5800-3.X5000-0.
5. Gunes I. Examination of wind effect on ADSS cables aging test / I. Gunes // Electrica. – 2018. – Vol. 18, No. 2. – P. 321 – 324. DOI: 10.5152/iujeee.2018.1818.
6. Tingting Zhan. Optimization of the optical fiber with triple sector stress elements / Zhan Tingting, Ji Minning // Optical Fiber Technology. – 2020. – Vol. 57. – Article no. 102212. DOI: 10.1016/j.yofte.2020.102212.
7. IEEE Guide for Overhead Alternating Current (AC) Transmission Line Design. – IEEE Standard 1863, 2024. DOI: 10.1109/IEEESTD.2024.10622047.
8. Рекомендації з підвішування оптичних кабелів на опорах повітряних ліній зв'язку, ЛЕП, контактної мережі залізниць: Р 45-010-2002. – [Чинний від 27.04.2003 р.]. – К. : Державний комітет зв'язку та інформатизації України, 2004. – 95 с.
9. Canudo Jan. Overhead Transmission Line Sag Monitoring Using a Chirped-Pulse Phase-Sensitive OTDR / [Canudo Jan., P. Sevillano, A. Iranzo et al.] // IEEE Sensors Journal, 2024, January. – Vol. 24, № 2. – P. 1988–1995. DOI: 10.1109/JSEN.2023.3340296.
10. The external mechanical effects on the value of additional losses in the telecommunications fiber optic cables under operating conditions / [A. Mekhtiyev, F. Bulatbayev, Y. Neshina et al.] // Elektrotechnik, Maschinenbau und Wirtschaftsingenieurwesen. – 2018. – Vol. 6. – P. 123–127. DOI:10.13142/kt10006.43.
11. Лінійно-кабельні споруди телекомунікацій. Проектування : ГБН В.2.2-34620942-002:2015. – [Чинний від 01.08.2015 р.]. – К. : Адміністрація Держспецзв'язку, 2015. – 140 с.
12. Determination of the coefficient of thermal expansion with embedded long-gauge fiber optic sensors / [X. Feng, Changsen Sun, Xiaotan Zhang et al.] // Measurement Science and Technology. – 2010. – Vol. 21, No. 6. – Article no. 065302. DOI:10.1088/0957-0233/21/6/065302.


DOI 10.15588/1607-3274-2024-4-2

УДК 621.396

# ОПТИМІЗАЦІЯ СТРУКТУР РАДІОСИСТЕМ ПЕЛЕНГАЦІЇ ДЖЕРЕЛ ВИПРОМІНЮВАННЯ СИГНАЛІВ З ПОВНІСТЮ ВІДОМИМИ ПАРАМЕТРАМИ

**Жила С. С.** – д-р техн. наук, доцент, завідувач кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Церне Е. О.** – д-р філософії, молодший науковий співробітник кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Попов А. В.** – д-р техн. наук, доцент, професор кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Руженцев М. В.** – д-р техн. наук, професор, головний науковий співробітник кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Волков Є. Г.** – директор, Державне підприємство «Науково-дослідний інститут “ОРІОН”», Київ, Україна.

**Шевчук С. Д.** – заст. нач. від. № 91, Державне підприємство «Науково-дослідний інститут “ОРІОН”», Київ, Україна.

**Грибський О. П.** – старш. наук. співроб., Державне підприємство «Науково-дослідний інститут “ОРІОН”», Київ, Україна.

**Колесніков Д. В.** – д-р філософії, старший викладач кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Інкарбаєва О. С.** – д-р філософії, старший викладач кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Черепнін Г. С.** – д-р філософії, старший викладач кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

## АНОТАЦІЯ

**Актуальність.** Радіопеленгатори є ключовими компонентами систем радіолокації та радіонавігації, особливо коли вони встановлюються на борту БПЛА. Високі вимоги до точності пеленгації та широкого кута однозначних вимірювань стають особливо актуальними в умовах збільшення застосування безпілотних систем. Основна проблема полягає у досягненні балансу між високою точністю та широким діапазоном кутів однозначних вимірювань.

**Мета.** Одночасне підвищення точності пеленгації та розширення кутів однозначних вимірювань за рахунок статистичного синтезу методів обробки функціонально-детермінованих сигналів в багатоканальних радіопеленгаторах.

**Метод.** Ґрунтується на статистичній теорії оптимізації радіотехнічних систем дистанційного зондування та радіолокації. Для обмеженого в даній роботі типу сигналів, що задаються функціонально-детермінованими моделями, сконструйовано функцію правдоподібності та визначено її максимуми для різних конфігурацій багатоантенних пеленгаторів. Результати статистичного синтезу перевіряються методами імітаційного моделювання та натурними експериментами.

**Результати.** Теоретичними дослідженнями та імітаційним моделюванням підтверджено, що в двоантених радіопеленгаторах існує протиріччя між високою роздільною здатністю та шириною діапазону кутів однозначного радіопеленгування. Отримано удосконалений метод обробки сигналів в чотирьохантенному радіопеленгаторі, що має пару високоспрямованих та пару низькоспрямованих антен. Для отримання гранично досяжної точності пеленгування у межах діапазону однозначних вимірювань радіопеленгатору синтезовано новий метод обробки сигналів в шести елементному радіоприймачі, що комплексує обробку в двох амплітудних пеленгаторах та одному фазовому пеленгаторі.

**Висновки.** Запропонований підхід дозволяє досягти оптимального балансу між роздільною здатністю і діапазоном кутів, що особливо важливо для застосування у бортових системах безпілотних літальних апаратів. Результати моделювання підтверджують ефективність запропонованого методу, що робить його перспективним для впровадження в сучасні радіосистеми.

**КЛЮЧОВІ СЛОВА:** багатоантенні пеленгатори, статистична оптимізація, оптимальне оброблення сигналів, імітаційне моделювання.

## АБРЕВІАТУРИ

БПЛА – безпілотний літальний апарат.

## НОМЕНКЛАТУРА

$\dot{s}_i(t,\theta_s)$ – прийняті радіопеленгатором корисні сигнали;


DOI 10.15588/1607-3274-2024-4-3

$K_{0i}$ – коефіцієнт підсилення $i$-го приймального каналу;

$\dot{G}_i(\theta-\theta_{0i})$ – діаграма спрямованості $i$-ї антени, що орієнтована своїм максимумом у напрямку $\theta_{0i}$;

$\theta_{0i}$ – напрям максимуму діаграми спрямованості $i$-ї антени;

$\delta(\theta-\theta_s)$ – дельта-функція, що визначає просторове положення точкового джерела радіовипромінювання у напрямку $\theta_s$;

$\delta(t_1-t_2)$ – дельта-функція у часі;

$\dot{A}(t)$ – комплексна обвідна сигналу, що випромінюється джерелом;

$j$ – мнима одиниця;

$f_0$ – несуча частота;

$t$ – час;

$\psi_i(\theta_s)$ – фазовий зсув сигналу в кожному приймальному каналі відносно фазового центру поля антен;

$\theta$ – координати кутів;

$c$ – швидкість розповсюдження електромагнітних хвиль;

$\Delta r_i(\theta_s)$ – різниця відстаней, що проходять електромагнітні хвилі від джерела до кожної антени;

$n_i(t)$ – внутрішні шуми в приймальних трактах;

$0,5N_0$ – спектральну щільність потужності внутрішніх шумів;

$R_{n_i}(t_1-t_2)$ – кореляційна функція внутрішніх шумів;

$\vec{u}(t)$ – вектор рівнянь спостереження;

$\mathrm{Re}\{\cdot\}$ – оператор дійсної частини комплексного сигналу;

$P[\vec{u}(t)\,|\,\lambda]$ – умовний функціонал щільності ймовірності випадкового процесу $\vec{u}(t)$;

$\lambda$ – параметр, що підлягає оцінюванню;

$\lambda_{true}$ – істинне значення параметра $\lambda$;

$\kappa$ – коефіцієнт в $P[\vec{u}(t)\,|\,\lambda]$, що не залежить від параметрів, що підлягають оцінюванню;

$T$ – час спостереження;

$\theta_s$ – кутове положення джерела випромінювання корисного сигналу;

$\theta_b$ – рівносигнальний напрямок;

$\theta_\delta$ – кут на який рознесені діаграми амплітудного пеленгатора;

$k_{\theta_i}$ – крутизна нормованих діаграм спрямованості $i$-ї антени;

$\dot{h}(t)$ – оптимальний фільтр;

$Y(\theta_s)$ – оптимальний вихідний ефект радіопеленгатору;

$\Psi(\cdot)$ – пеленгаційна характеристика радіопеленгатору;

$E_s$ – енергія сигналу, що випромінює джерело;

$d$ – відстань між антенами фазового пеленгатору

$a_i$ – амплітудний множник в моделі діаграми спрямованості;

$\Delta\theta$ – ширина діаграми спрямованості;

$\sigma^2_\lambda$ – гранична дисперсія оцінки $\lambda$;

$\mu_0$ – співвідношення сигнал-завада.

## ВСТУП

Радіопеленгатори є важливою складовою майже всіх сучасних систем і комплексів радіолокації [1], радіомоніторингу [2], радіонавігації [3], радіонаведення [4] і більшості випадків визначають їх технічні характеристики. Діапазон робочих частот, точність пеленгації, габаритні розміри і діапазон кутів однозначних вимірювань є основними характеристиками пеленгаторів [5, 6]. Особливо високі вимоги до цих характеристик та систем в цілому висувають, коли пеленгатор встановлюється на борт БПЛА. Сучасні тенденції [7–9] збільшення галузей застосування безпілотних літальних апаратів, безекіпажних морських суден та наземних транспортних засобів вимагають підвищення якості роботи бортових систем визначення кутового положення джерел радіовипромінювання. При цьому найбільш актуальними є питання вирішення основного протиріччя в роботі радіопеленгатору між високою точністю пеленгації та широким кутом однозначних вимірювань.

**Об’єктом дослідження** є процес вимірювання кутового положення джерел радіовипромінювання багатоантенними радіосистемами.

**Предмет дослідження** – оптимальні методи та алгоритми обробки сигналів амплітудних та фазових радіопеленгаторах.

**Мета дослідження** полягає в одночасному підвищенні точності пеленгації та розширенні кутів однозначних вимірювань за рахунок статистичного синтезу методів обробки функціонально-детермінованих сигналів в багатоканальних радіопеленгаторах.

## 1 ПОСТАНОВКА ЗАДАЧІ

З актуальності задачі випливає проблема статистичного синтезу та аналізу методів високоточного та однозначного визначення кутових положень джерел випромінювання функціонально-детермінованих сигналів в багатоантенних радіопеленгаторах.

Для вирішення проблеми необхідно вирішити часткові завдання:

1) визначення моделей прийнятих радіопеленгатором сигналів, шумів та їх статистичних характеристик,

2) розвиток теоретичних основ статистичної оптимізації методів вимірювання кутових положень дже-


DOI 10.15588/1607-3274-2024-4-3

рел радіовипромінювання в багатоантенних радіопеленгаторах,

3) вирішення часткових задач оптимізації структур радіопеленгаторів з різною конфігурацією,

4) аналіз точності та діапазону кутів однозначних вимірювань в отриманих методах.

Нехай багатоантенний радіопеленгатор приймає сигнали, що описуються загальною функціонально-детермінованою моделлю

$$\dot{s}_i(t,\theta_s) = \\ = K_{0i}\int_{\Theta}\dot{G}_i(\theta-\theta_{0i})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}e^{j\psi_i(\theta_s)}d\theta, \quad (1)$$

$\psi_i(\theta_s) = 2\pi f_0 \Delta r_i(\theta_s)c^{-1}$, $\Delta r_i(\theta_s) = x_i\cos\theta_s + y_i\sin\theta_s$, $e^{-j2\pi f_0 t}$ – гармонійне коливання у комплексній формі з несучою частотою $f_0$, $i=\overline{1,N}$. Геометрія, що лягла в основу моделі (1), показана на рис. 1.

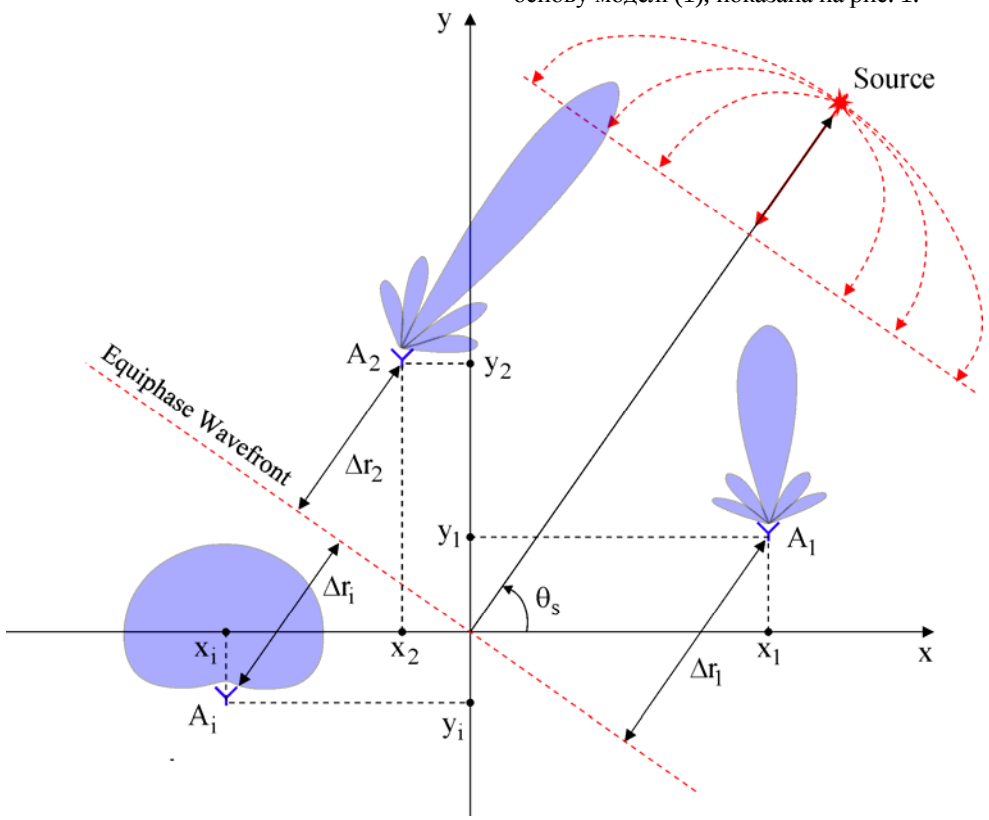



Рисунок 1 – Загальна геометрія вимірювання кутового положення джерела радіовипромінювання Source в багатоантенному пеленгаторі (Equiphase Wavefront рівнофазний хвильовий фронт)

Також визначимо обмеження, що накладаються на апріорні відомості про вплив шумів на прийняті сигнали. Внутрішні шуми в приймальних трактах $n_i(t)$ будемо вважати білими гаусівськими шумами, що взаємно не корельовані, проте мають однакову спектральну щільність потужності $0,5N_0$. Кореляційна функція внутрішніх шумів має вигляд

$$R_{n_i}(t_1-t_2) = 0,5N_0\delta(t_1-t_2). \quad (2)$$

Модель зв'язку між сигналами та шумами визначимо у вигляді їх адитивної суміші

$$\vec{u}(t) = \mathrm{Re}\{\dot{\vec{s}}(t,\theta_s)\} + \vec{n}(t), \quad (3)$$

$$\vec{u}(t) = \left\|u_1(t),\, u_2(t),\, ...,\, u_N(t)\right\|, \quad (4)$$

$$\dot{\vec{s}}(t,\theta_s) = \left\|\dot{s}_1(t,\theta_s),\, \dot{s}_2(t,\theta_s),\, ...,\, \dot{s}_N(t,\theta_s)\right\|, \quad (5)$$

$$\vec{n}(t) = \left\|n_1(t),\, n_2(t),\, ...,\, n_N(t)\right\|. \quad (6)$$

В таких припущення задача вимірювання кутового положення $\theta_s$ джерела радіовипромінювання полягає в необхідності синтезувати оптимальний, найкращій у межах певного критерію, метод обробки прийнятих корисних сигналів на фоні внутрішніх шумів.

## 2 ОГЛЯД ЛІТЕРАТУРИ

Існує декілька базових принципів радіопеленгації, що ґрунтуються на вимірюванні параметрів та статистичних характеристик прийнятих сигналів. Основні типи пеленгаторів – амплітудні, фазові, кореляційні.

В амплітудних пеленгаторах [10] бувають наступні варіанти реалізації вимірювань: обертання однієї високоспрямованої антени, перемикання між декількома високоспрямованими антенами, електронна комутація низькоспрямованих антен в круговій антенній решітці


DOI 10.15588/1607-3274-2024-4-3

створюючи ефект обертання вузького променю по колу, проведення вимірювань парою антен Адкока. Останні розробки методів та систем показують підвищення точності амплітудної пеленгації джерел радіовимірювань, проте були розроблені без обмежень на габаритні розміри носія радіопеленгатору.

Фазові пеленгатори [11] реалізуються за наступними принципами: безпосереднє вимірювання різниці фаз у приймачах з антенами, що розміщенні на відстанях менших за половину довжини хвилі, інфремерометричне вимірювання різниці фаз в декількох просторово-роззосереджених приймачах, вимірювання фази прийнятих коливань з фазовою модуляцією, що створена ефектом Доплера при круговому обертанні (чи створені ефекту такого обертання) приймальної антени. Аналіз наукових розробок та ринку продукції таких пеленгаторів показує значне підвищення тості пеленгації, удосконалення цифрових методів обробки сигналів та зменшення їх габаритних розмірів. В той самий час вимірювання фази у будь-яких радіосистемах дозволяє вирішувати задачі в дуже вузькому діапазоні значень або з неоднозначністю з подальшим її усуненням вторинною обробко. Дане протиріччя так і не вирішено в сучасній літературі.

Кореляційні пеленгатори [12] об'єднують переваги амплітудних та фазових пеленгаторів, визначаючи напрямок на джерело радіовипромінювання методами статистичної обробки комплексних амплітуд прийнятих сигналів. За наявності достатньої бази антенної решітки в таких пристроях можливе роздільне пеленгування сигналу і завади в суміщеному каналі та виділення сигналу на тлі когерентних завад. У них застосовуються так звані алгоритми множинної класифікації сигналу (MUSIC), багаторівневі методи оцінки максимальної правдоподібності (MLE), а також алгоритми надроздільної здатності (SR-DF). Кореляційні пеленгатори мають найкращі характеристики, проте потребують значних апріорних відомостей. Для прикладу реалізувати повний потенціал методу MUSIC можливо лише при обробці функціонально-детермінованих сигналів, коли відома частота випромінювання і кількість джерел радіовипромінювання. Також кореляційні методи складні за реалізацією та мають високу ціну виробництва через необхідність обробляти комплексні амплітуди когерентно.

## 3 МАТЕРІАЛИ ТА МЕТОДИ

Для розв'язання поставленої задачі синтезу методу обробки сигналів в багатоканальних радіопеленгаторах будемо використовувати статистичну теорію оптимізації радіотехнічних систем дистанційного зондування та радіолокації, що розвинута в роботах професора Валерія Волосюка [13, 14]. Згідно цієї теорії для вирішення зазначеної проблеми доцільно використовувати метод максимуму функції правдоподібності. Сутність цього методу полягає в пошуку параметру $\lambda$, що максимізує функціонал правдоподібності $P[\vec{u}(t)\,|\,\lambda]$ – умовний функціонал щільності ймовірності випадкового процесу $\vec{u}(t)$ при фіксованому значенні параметра $\lambda$. Замість функціоналу $P[\vec{u}(t)\,|\,\lambda]$ частіше максимізують його логарифм, так як функція логарифму є монотонною і не змінює точку максимуму $P[\vec{u}(t)\,|\,\lambda]$. Для знаходження оптимальних оцінок параметра $\lambda$ необхідно розв'язати систему рівнянь

$$\left.\frac{d\ln P[\vec{u}(t)\,|\,\lambda]}{d\lambda}\right|_{\lambda=\lambda_{true}} = 0, \qquad (7)$$

де $d/d\lambda$ – оператор похідної, яка береться в точці істинного значення $\lambda_{true}$ параметра $\lambda$.

Одним із найважливіших етапів розв'язання оптимізаційної задачі є визначення $P[\vec{u}(t)\,|\,\lambda]$. В [15] наведена методика конструювання функціоналів правдоподібності для широкого кола задач радіолокації та дистанційного зондування. В цьому дослідженні будемо використовувати функціонал правдоподібності наступного вигляду:

$$P[\vec{u}(t)\,|\,\lambda=\theta_s] = \\ = \kappa\exp\left\{-\frac{1}{N_0}\sum_{i=1}^{N}\int_T\left[u_i(t)-\mathrm{Re}\left\{\dot{s}_i(t,\theta_s)\right\}\right]^2 dt\right\}. \qquad (8)$$

Параметр $\lambda=\theta_s$ є постійною величиною.

Підставляючи (8) в (7), отримаємо систему рівнянь правдоподібності

$$\frac{2}{N_0}\sum_{i=1}^{N}\int_T\left[u_i(t)-\mathrm{Re}\left\{\dot{s}_i(t,\theta_s)\right\}\right]\mathrm{Re}\left\{\frac{d\dot{s}_i(t,\theta_s)}{d\theta_s}\right\}dt=0 \qquad (9)$$

або

$$\sum_{i=1}^{N}\int_T u_i(t)\,\mathrm{Re}\left\{\frac{d\dot{s}_i(t,\theta_s)}{d\theta_s}\right\}dt = \\ = \sum_{i=1}^{N}\int_T \mathrm{Re}\left\{\dot{s}_i(t,\theta_s)\right\}\mathrm{Re}\left\{\frac{d\dot{s}_i(t,\theta_s)}{d\theta_s}\right\}dt. \qquad (10)$$

Ліва частина (10) визначає основні оптимальні операції, що необхідно виконати над прийнятими коливаннями $u_i(t)$ в кожному каналі. Ліву частину необхідно порівняти з правою частиною. Права частина – це пеленгаційна характеристика багатоканального радіопеленгатору, що показує реакцію вимірювача на зміну кутового положення джерела випромінювання $\theta_s$. На першому етапі визначають аналітичний вигляд пеленгаційної характеристики на основі заданих параметрів сигналів, що випромінює джерело, з урахування просторових характеристик антен, геометрії їх розміщення, часу спостереження та параметрів вхід-


DOI 10.15588/1607-3274-2024-4-3

них трактів. В подальшому отриману пеленгаційну характеристику необхідно практично підтвердити експериментом і в реалізованому пристрої підставляти результати вимірювань. Відмінність від теорії обумовлена багатьма аспектами функціонування кожного окремого елемента приймача, відхиленням його параметрів від номінальних, впливом флікер-шуму, різними показниками узгодженості НВЧ елементів тощо.

Конкретизуючи корисні сигнали в рівнянні спостереження, на основі виразу (10) можливо отримати вирішення часткових оптимізаційних задач синтезу алгоритмів в різних типах радіопеленгаторів.

Для синтезу методу радіопеленгації, що подолав би протиріччя між високою точністю визначення кута та широким діапазоном однозначних вимірювань доцільно спочатку вирішити часткову задачу оптимізації структури двоканального амплітудного пеленгатору і показати його недоліки. Для усунення недоліків пропонується синтезувати чотирьохантенний амплітудний пеленгатор. Також приймаючи до уваги, що найвищу точність мають фазові пеленгатори, важливо розробити алгоритм комплексування вимірювань в амплітудних і фазових пеленгаторах. Далі будуть вирішені ці часткові завдання.

Синтезуємо метод обробки сигналів в двоантенному амплітудному пеленгаторі. Будемо вважати, що сигнали від джерела випромінювання приймаються двома антенами, що мають ідентичні, але рознесені у просторі на кут $\theta_\delta$ діаграми спрямованості, $\dot{G}_1(\theta)=\dot{G}_2(\theta)$. Напрямок з якого в обох приймальних каналах будуть спостерігатися рівні значення сигналів називається рівносигнальним і позначено $\theta_b$. Типова геометрія опису вимірювань кута $\theta_s$ в двоантеному пеленгаторі показана на рис. 2. При визначені кута орієнтації максимуму діаграм спрямованості антен в (1) прийнято до уваги, що $\theta_{01}=(\theta_b+0,5\theta_\delta)$, $\theta_{02}=(\theta_b-0,5\theta_\delta)$.

Моделі корисних сигналів для зазначених умов вимірювання в двоантенному пеленгаторі набувають вигляд:

$$\dot{s}_1(t,\theta_s)= \\ =\int_\Theta \dot{G}(\theta-\theta_b-0,5\theta_\delta)\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta, \quad (11)$$

$$\dot{s}_2(t,\theta_s)= \\ =\int_\Theta \dot{G}(\theta-\theta_b+0,5\theta_\delta)\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta. \quad (12)$$

В наведених моделях (11), (12) відсутня інформація про фазовий зсув $\psi_i(\theta_s)$, тому що передбачається що обробка ведеться в кожному каналі некогерентно.

Рівносигнальний напрямок $\theta_b$ в двоантенних пеленгаторах зазвичай обирається на лінійних ділянках першої та другої діаграм і всі вимірювання також виконуються на лінійних ділянках. Розглянемо для прикладу діаграми спрямованості двох антен в полярній системі координат у вигляді двох дійсних гаусівських функцій, рис. 3. У межах лінійних ділянок $G_1(\theta)$ і $G_2(\theta)$ можуть бути розкладені в ряд Тейлора в околиці напрямку $\theta_b$ і обмежені лінійним наближенням

$$G(\theta-\theta_b-0,5\theta_\delta)=G(\theta_b)(1+k_{\theta 1}(\theta-\theta_b)), \\ G(\theta-\theta_b+0,5\theta_\delta)=G(\theta_b)(1+k_{\theta 2}(\theta-\theta_b)), \quad (13)$$

де

$$k_{\theta_1}=\left.\frac{dG(\theta-\theta_b-0,5\theta_\delta)}{d\theta}\right|_{\theta=\theta_b}, \\ k_{\theta_2}=\left.\frac{dG(\theta-\theta_b+0,5\theta_\delta)}{d\theta}\right|_{\theta=\theta_b}. \quad (14)$$

Для однакових і симетричних відносно $\theta_b$ діаграм $k_{\theta_2}=-k_{\theta_1}$. В подальшому будемо використовувати лише $k_\theta$ з різними знаками.

Приймаючи до уваги (13) рівняння спостереження приймуть вигляд

$$u_1(t)=\mathrm{Re}\left\{G(\theta_b)\int_\Theta(1-k_\theta(\theta-\theta_b))\times\right. \\ \left.\times\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}\,d\theta\right\}+n_1(t), \quad (15)$$

$$u_2(t)=\mathrm{Re}\left\{G(\theta_b)\int_\Theta(1+k_\theta(\theta-\theta_b))\times\right. \\ \left.\times\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}\,d\theta\right\}+n_2(t). \quad (16)$$

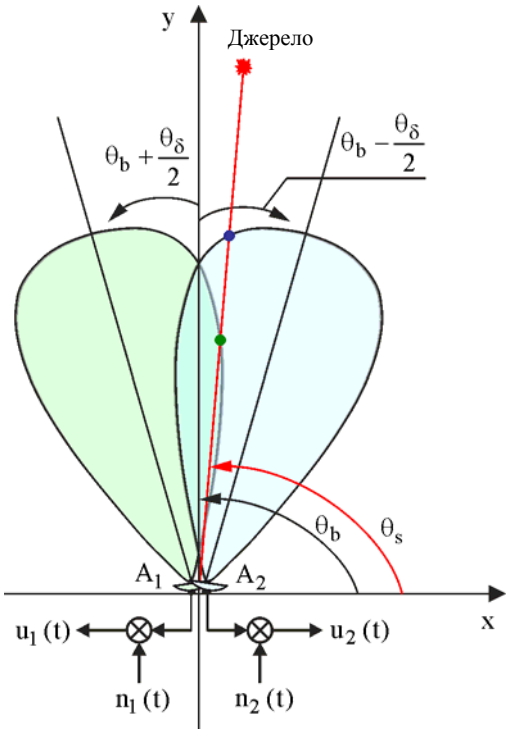



Рисунок 2 – Геометрія вимірювань в двоантенному пеленгаторі


DOI 10.15588/1607-3274-2024-4-3

Підставляючи (11), (12), (15), (16) в (10) отримуємо нерівність

$$\int_T u_1(t)\dot{A}(t)e^{-j2\pi f_0 t}dt - \int_T u_2(t)\dot{A}(t)e^{-j2\pi f_0 t}\ dt =$$

$$= \frac{1}{2}\int_T \dot{s}_1(t,\theta_s)\dot{A}^*(t)e^{j2\pi f_0 t}\,dt -$$

$$-\frac{1}{2}\int_T \dot{s}_2(t,\theta_s)\dot{A}^*(t)e^{j2\pi f_0 t}dt\,. \qquad (17)$$

Вводячи коефіцієнт передачі оптимального фільтру $\dot{h}(t)=\dot{A}(t)e^{-j2\pi f_0 t}$ отримуємо в лівій частині нерівності (17) оптимальний вихідний ефект радіопеленгатору

$$Y(\theta_s)=\int_T u_2(t)\dot{h}(t)dt - \int_T u_1(t)\ \dot{h}(t)\ dt\,. \qquad (18)$$

Права частина (17) є пеленгаційною характеристикою, що набуває вигляду лінійної функції після підстановки (11) і (12)

$$\Psi(\theta_s)=\chi(\theta_s-\theta_b)\,, \qquad (19)$$

де $\chi = k_\theta \int_T |\dot{A}(t)|^2\,dt$.

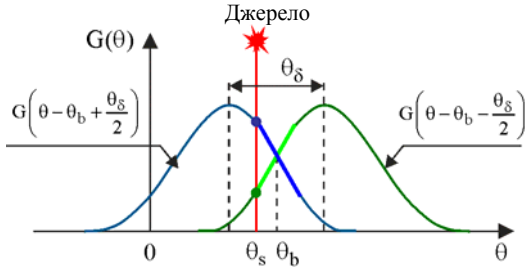



Рисунок 3 – Діаграми спрямованості в двоантенному пеленгаторі

Для введеного у даній частковій задачі спрощення щодо розрахунків у межах лінійної ділянки діаграм спрямованості можливо отримати алгоритм вимірювання $\theta_s$, підставивши (19) і (18) в (17)

$$\theta_s = \frac{Y(\theta_s)}{\chi} + \theta_b\,. \qquad (20)$$

Основними математичними операціями обробки сигналів в двоантенному амплітудному радіопеленгаторі є узгоджена фільтрація в фільтрі $\dot{h}(t)$ та операція віднімання результатів фільтрації в різних каналах. Коефіцієнт $\chi$ попереднього практичного вимірювання для реалізації алгоритму (19).

Діапазон однозначних вимірювань в отриманому алгоритмі визначається розміром лінійної ділянки, а точність пропорційна крутизні $k_\theta$. Зі збільшенням крутизни отримуємо вищу точність пеленгації, але при цьому звужується діапазон кутів однозначного вимірювання положення джерела радіовипромінювання.

Синтезуємо метод обробки сигналів в чотирьохантенному амплітудному пеленгаторі. Для подолання протиріччя між точністю та однозначністю вимірювань у широкому діапазоні кутів пропонується використовувати дві пари антен. Перша пара антен має широкі діаграми спрямованості, а друга пара – вузькі. Таким чином при комплексування вимірювань можливе об'єднання переваг двох радіопеленгаторів.

Геометрія проведення вимірювань чотирьохантенним радіопеленгатором показана на рис. 4.

Рівняння спостереження за геометрією на рис. 4 має вигляд

$$u_1(t) =$$

$$= \mathrm{Re}\{\int_\Theta G_1(\theta-\theta_b-\frac{\theta_{\delta 1}}{2})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta\} + n_1(t),$$

$$u_2(t) =$$

$$= \mathrm{Re}\{\int_\Theta G_1(\theta-\theta_b+\frac{\theta_{\delta 1}}{2})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta\} + n_2(t),$$

$$u_3(t) =$$

$$= \mathrm{Re}\{\int_\Theta G_2(\theta-\theta_b-\frac{\theta_{\delta 2}}{2})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta\} + n_3(t),$$

$$u_4(t) =$$

$$= \mathrm{Re}\{\int_\Theta G_2(\theta-\theta_b+\frac{\theta_{\delta 2}}{2})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta\} + n_4(t). \qquad (21)$$

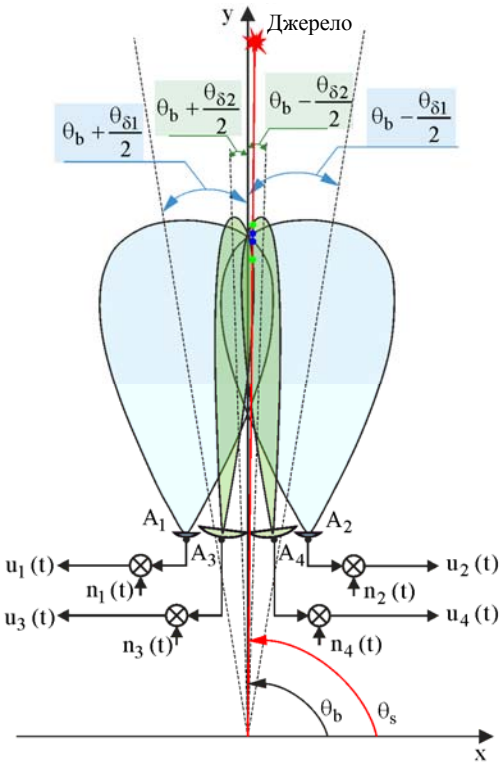



Рисунок 4 – Геометрія вимірювань в чотирьохантенному пеленгаторі


DOI 10.15588/1607-3274-2024-4-3

Підставляючи рівняння спостереження (20) і корисні сигнали, що в нього входять, в (9), отримуємо рівняння правдоподібності

$$\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}\int_T u_1(t)\dot{h}(t)\,dt+$$
$$+\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}\int_T u_2(t)\dot{h}(t)\,dt+$$
$$+\left.\frac{dG_2(\theta-\theta_b-0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}\int_T u_3(t)\dot{h}(t)\,dt+$$
$$+\left.\frac{dG_2(\theta-\theta_b+0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}\int_T u_4(t)\dot{h}(t)\,dt=$$
$$=\frac{1}{2}E_s\,\Psi(\theta_s,\theta_b,\theta_{\delta 1},\theta_{\delta 2}), \qquad (22)$$

де $E_s=\int_T\left|\dot{A}(t)\right|^2dt$,

$$\Psi(\theta_s,\theta_b,\theta_{\delta 1},\theta_{\delta 2})=$$
$$=\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}G_1(\theta_s-\theta_b-0,5\theta_{\delta 1})+$$
$$+\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}G_1(\theta_s-\theta_b+0,5\theta_{\delta 1})+ \qquad (23)$$
$$+\left.\frac{dG_2(\theta-\theta_b-0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}G_2(\theta_s-\theta_b-0,5\theta_{\delta 2})+$$
$$+\left.\frac{dG_2(\theta-\theta_b+0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}G_2(\theta_s-\theta_b+0,5\theta_{\delta 2})$$

– пеленгаційна характеристика чотирьохантенної радіолокаційної системи вимірювання кутового положення джерел радіовипромінювання.

Отриманий алгоритм доволі складний з технічної точки зору та практичних особливостей його реалізації, але основні математичні операції зрозумілі.

На відміну від пеленгаційної характеристики у двоантенному пеленгаторі (18), що має вигляд лінійної функції, в цих розрахунках ми вирішили більш загальну задачу аналізу роботи пеленгатору. По-перше, в отриманій формулі для $\Psi(\cdot)$ описує будь-який вигляд діаграми спрямованості, що може відрізнятися для пар антен. По-друге, ми не обмежувались лише лінійною ділянкою роботи радіопеленгатору і отримали загальні множники $\frac{dG_i(\cdot)}{d\theta}$ швидкості зміни функцій $G_i(\cdot)$ в кожному каналі. Це дозволить проаналізувати весь можливий діапазон кутів огляду антенами і обрати найбільш широкий з однозначними вимірюваннями.

Ліва частина (21) є звичайним ваговим додаванням результату вимірювання і не показує в явному вигляді відомі алгоритмічні операції. В той самий час $\frac{dG_i(\cdot)}{d\theta}$ вагові коефіцієнти матимуть різний знак поблизу рівно сигнального напрямку і призведуть до відомих алгоритмів віднімання результатів вимірювань в різних каналах.

Більш детальний аналіз отриманих виразів для випадку функціонального опису діаграм спрямованості буде наведено далі при імітаційному моделюванні пеленгаційних характеристик.

Отримані алгоритми об'єднують переваги двоантенних пеленгаторів з вузькими та широкими діаграмами, дозволяють подолати протиріччя між точністю та діапазоном кутів однозначних вимірювань. В той самий час представляє інтерес отримати алгоритм з найвищою точністю пеленгації без втрати сигналу у широкому діапазоні кутів.

Синтезуємо метод обробки сигналів в шестиантенному амплітудному пеленгаторі. Для досягнення найвищої точності вимірювання кутів пропонується додати до чотирьохантенного пеленгатору два канали вимірювання різниці фаз сигналів. В такій схемі вимірювання антени фазового пеленгатору можуть бути неспрямовані, проте в прийнятих сигналах необхідно врахувати зсув фаз відносно фазового центру. Геометрія вимірювань показана на рис. 5.

Шуми в кожному каналі також вважаємо гаусівськими білими випадковими процесами, що некорельовані за часом та між каналами

$$R_{n_1}(t_1-t_2)=R_{n_2}(t_1-t_2)=R_{n_3}(t_1-t_2)=$$
$$=R_{n_4}(t_1-t_2)=R_{n_5}(t_1-t_2)=R_{n_6}(t_1-t_2)=$$
$$R_{n_1}(t_1-t_2)=R_{n_2}(t_1-t_2)=R_{n_3}(t_1-t_2)=$$
$$=R_{n_4}(t_1-t_2)=R_{n_5}(t_1-t_2)=R_{n_6}(t_1-t_2)= \qquad (24)$$
$$=0,5N_0\delta(t_1-t_2).$$

Корисні сигнали мають наступні моделі:

$$\dot{s}_1(t,\theta_s)=$$
$$=K_{01}\int_\Theta G_1(\theta-\theta_b-0,5\theta_{\delta 1})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta,$$
$$\dot{s}_2(t,\theta_s)=$$
$$=K_{02}\int_\Theta G_1(\theta-\theta_b+0,5\theta_{\delta 1})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta,$$
$$\dot{s}_3(t,\theta_s)=$$
$$=K_{03}\int_\Theta G_2(\theta-\theta_b-0,5\theta_{\delta 2})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta,$$


DOI 10.15588/1607-3274-2024-4-3

$$\dot{s}_4(t,\theta_s)=$$
$$=K_{04}\int_{\Theta}G_2(\theta-\theta_b+0,5\theta_{\delta 2})\delta(\theta-\theta_s)\dot{A}(t)e^{-j2\pi f_0 t}d\theta,$$

$$\dot{s}_5(t,\theta_s)=K_{05}G_3(\theta_b)\dot{A}(t)e^{-j2\pi f_0\left(t-\frac{1}{2}\frac{d\cos(\theta_s-\theta_b)}{c}\right)},$$

$$\dot{s}_6(t,\theta_s)=K_{06}G_3(\theta_b)\dot{A}(t)e^{-j2\pi f_0\left(t+\frac{1}{2}\frac{d\cos(\theta_s-\theta_b)}{c}\right)}. \quad (25)$$

Підставляючи рівняння спостереження, що складаються з адитивної суміші корисних сигналів та дельта-корельованих шумів, в нерівність (10), отримуємо

$$\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}K_{01}\int_T u_1(t)\dot{h}(t)\,dt+$$
$$+\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}K_{02}\int_T u_2(t)\dot{h}(t)\,dt\tilde{+}$$
$$+\left.\frac{dG_2(\theta-\theta_b-0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}K_{03}\int_T u_3(t)\dot{h}(t)\;dt+$$
$$+\left.\frac{dG_2(\theta-\theta_b+0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}K_{04}\int_T u_4(t)\dot{h}(t)\,dt-$$
$$-G_3(\theta_b)\sin(\theta_s-\theta_b)\times$$
$$\times\left(\cos\left(\pi f_0\frac{d\cos(\theta_s-\theta_b)}{c}\right)+j\sin\left(\pi f_0\frac{d\cos(\theta_s-\theta_b)}{c}\right)\right)\times$$
$$\times\left(j\pi f_0 dc^{-1}\right)K_{05}\int_T u_5(t)\dot{h}(t)\,dt+$$
$$+G_3(\theta_b)\sin(\theta_s-\theta_b)\times$$
$$\times\left(\cos\left(\pi f_0\frac{1}{2}\frac{d\cos\theta_s}{c}(\theta_s-\theta_b)\right)-j\sin\left(\pi f_0\frac{d\cos(\theta_s-\theta_b)}{c}\right)\right)\times$$
$$\times(j\pi f_0 dc^{-1})K_{06}\int_T u_6(t)\dot{h}(t)dt=$$
$$=K_{01}\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}G_1(\theta_s-\theta_b+0,5\theta_{\delta 1})\frac{1}{2}E_s+$$
$$+K_{02}\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta}\right|_{\theta=\theta_s}G_1(\theta_s-\theta_b-0,5\theta_{\delta 1})\frac{1}{2}E_s+$$
$$+K_{03}\left.\frac{dG_2(\theta-\theta_b-0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}G_2(\theta_s-\theta_b+0,5\theta_{\delta 2})\frac{1}{2}E_s+$$
$$+K_{04}\left.\frac{dG_2(\theta-\theta_b+0,5\theta_{\delta 2})}{d\theta}\right|_{\theta=\theta_s}G_2(\theta_s-\theta_b-0,5\theta_{\delta 2})\frac{1}{2}E_s+$$
$$+(K_{06}-K_{05})G_3^2(\theta_b)\sin(\theta_s-\theta_b)\left(\pi f_0\frac{d}{c}\right)\frac{1}{2}E_s\,. \quad (26)$$

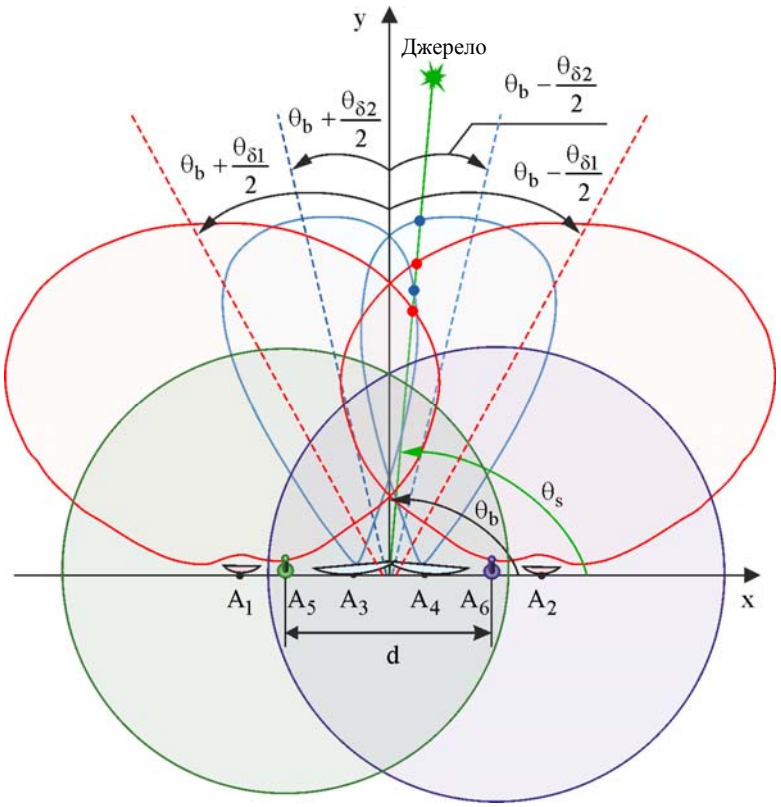



Рисунок 5 – Геометрія вимірювань в шестиантенному пеленгаторі


DOI 10.15588/1607-3274-2024-4-3

Як і раніше, ліва частина (26) показує оптимальні операції, що необхідно виконати на прийнятими сигналами $\vec{u}(t)$ в кожному приймальному каналі пеленгатору. Перші чотири рівняння повторюють операції, що необхідно виконати в чотирьохантенному пеленгаторі. П'ятий і шостий доданок доволяють отримати оцінку кутового положення джерела радіовипромінювання з точністю реалізації фазових вимірювань. Основні операції наступні: 1) рівняння $u_5(t)$ $u_6(t)$ необхідно пропустити через фільтр, параметри якого узгоджені з комплексною обвідною і частотою сигналу випромінювача, 2) результат детектування на низькій частоті множаться нормуючі на коефіцієнти $G_3(\theta_b)(j\pi f_0 dc^{-1})K_{06}$, 3) нормовані амплітуди необхідно обробити в квадратурному детекторі, що налаштований на вимірювання зсуву фази що виникає через відхилення джерела випромінювання від рівносигнальної зони. Результати обробки перераховуються в кути за допомогою пеленгаційної кривої, що розраховується в правій частині (26). При заданих параметрах розміщення та орієнтації антен в пеленгаторі права частина (26) є функцією кута $\theta_s$. Аналітично дати оцінку отриманим виразам або провести їх аналіз без конкретизації форм діаграм спрямованості неможливо. Доцільно виконати імітаційне моделювання пеленгаційних характеристик та розрахунок граничних похибок вимірювання.

## 4 ЕКСПЕРИМЕНТИ

Спочатку проаналізуємо вже отриманий результат – праві частини рівнянь правдоподібності (16), (21) і (24), що представляють собою пеленгаційні характеристики різних типів багатоантенних радіопеленгаторів. Отримане рівняння (18) є лише лінійна частина такої характеристики, що не залежить від виду діаграм спрямованості і їх орієнтації у просторі. Для наочного порівняння всіх отриманих методів пеленгаційну характеристику двоантенного пеленгатора отримаємо з аналізу пеленгаційної характеристики чотирьохантенного пеленгатора. Імітаційне моделювання буде виконане за наступними виразами:

1) для двохантенного амплітудного пеленгатора

$$\Psi(\theta_s)=\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta1})}{d\theta}\right|_{\theta=\theta_s} G_1(\theta_s-\theta_b-0,5\theta_{\delta1})+$$
$$+\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta1})}{d\theta}\right|_{\theta=\theta_s} G_1(\theta_s-\theta_b+0,5\theta_{\delta1}),$$

2) для чотирьохантенного амплітудного пеленгатора

$$\Psi(\theta_s)=\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta1})}{d\theta}\right|_{\theta=\theta_s} G_1(\theta_s-\theta_b-0,5\theta_{\delta1})+$$
$$+\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta1})}{d\theta}\right|_{\theta=\theta_s} G_1(\theta_s-\theta_b+0,5\theta_{\delta1})+$$
$$+\left.\frac{dG_2(\theta-\theta_b-0,5\theta_{\delta2})}{d\theta}\right|_{\theta=\theta_s} G_2(\theta_s-\theta_b-0,5\theta_{\delta2})+$$
$$+\left.\frac{dG_2(\theta-\theta_b+0,5\theta_{\delta2})}{d\theta}\right|_{\theta=\theta_s} G_2(\theta_s-\theta_b+0,5\theta_{\delta2}),$$

3) для шестиантенного амплітудно-фазового пеленгатора

$$\Psi(\theta_s)=K_{01}\left.\frac{dG_1(\theta-\theta_b-0,5\theta_{\delta1})}{d\theta}\right|_{\theta=\theta_s} G_1(\theta_s-\theta_b+0,5\theta_{\delta1})+$$
$$+K_{02}\left.\frac{dG_1(\theta-\theta_b+0,5\theta_{\delta1})}{d\theta}\right|_{\theta=\theta_s} G_1(\theta_s-\theta_b-0,5\theta_{\delta1})+$$
$$+K_{03}\left.\frac{dG_2(\theta-\theta_b-0,5\theta_{\delta2})}{d\theta}\right|_{\theta=\theta_s} G_2(\theta_s-\theta_b+0,5\theta_{\delta2})+$$
$$+K_{04}\left.\frac{dG_2(\theta-\theta_b+0,5\theta_{\delta2})}{d\theta}\right|_{\theta=\theta_s} G_2(\theta_s-\theta_b-0,5\theta_{\delta2})+$$
$$+(K_{06}-K_{05})G_3^2(\theta_b)\sin(\theta_s-\theta_b)\left(\pi f_0\frac{d}{c}\right).$$

Всі змінні в наведених виразах, окрім $\theta_s$, будуть константами. У якості функцій, що описують діаграми спрямованості $G_i(\cdot)$ розглянемо наступні:

1) функція Гауса

$$G_i(\theta)=a_i e^{-\frac{(\theta-\theta_{0i})^2}{2\Delta\theta^2}}, \qquad (27)$$

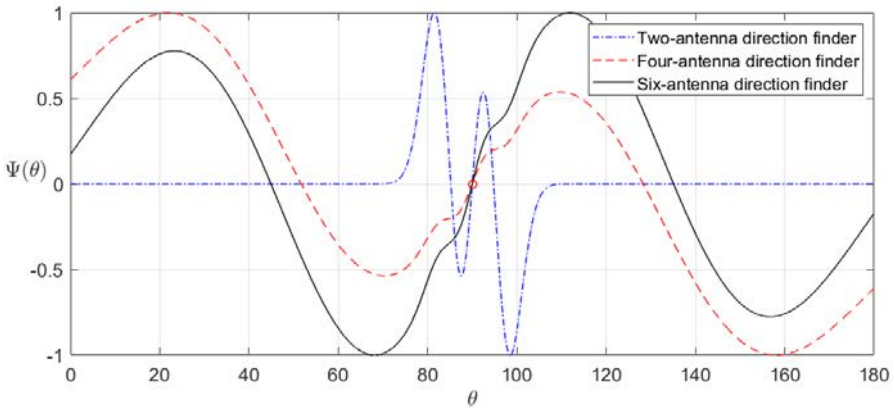
2) Sinc-функція

$$G_i(\theta)=a_i\,\mathrm{sinc}\left(\frac{\theta-\theta_{0i}}{\Delta\theta}\right). \qquad (28)$$

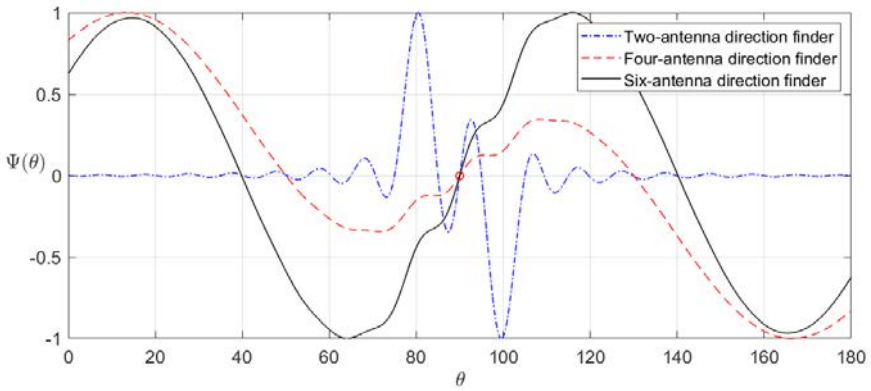
Нормовані до одиниці криві пеленгації джерел радіовипромінювання при використанні функції (27) з параметрами $\theta_b=90^\circ$, $\Delta\theta_1=30^\circ$, $\Delta\theta_2=5^\circ$, $K_{01}=K_{02}=K_{03}=K_{04}=(K_{06}-K_{05})=1$, $a_1=6$, $a_2=1$, $\theta_{\delta1}=60^\circ$, $\theta_{\delta2}=10^\circ$, $G_3(\theta_b)=1$, $f_0=1\,GHz$, $d=0,1\,m$ показані на рис. 6а.

Пеленгаційні характеристики для вже наведених параметрів при використанні антен з діаграмами (28) показані на рис. 6б.


DOI 10.15588/1607-3274-2024-4-3




а




б

Рисунок 6 – Нормовані до одиниці пеленгаційні криві (блакитний – двоантенний, червоний – чотирьоантенний, чорний – шести антенний пеленгатор): а – діаграми спрямованості у вигляді функцій Гауса, б − діаграми спрямованості у вигляді Sinc-функцій

З аналізу отриманих графіків слідує, що в двохантенному пеленгаторі дійсно існує протиріччя між діапазоном кутів однозначних вимірювань, що відповідають лінійній ділянці пеленгаційної характеристики, та реакцією приймача на зміну кута (точністю).

В чотирьох антенному пеленгаторі протиріччя вдалося подолати. З графіків випливає, що пеленгаційна характеристика має збільшений у 4 рази діапазон однозначних вимірювань, проте має дві ділянки з різною крутизною.

Центральна частина пеленгаційної характеристики має більшу крутизну і дозволяє в межах лінійної ділянки вузьких антен вирішувати задачу з високою точністю. За межами цієї ділянки крутизна менша, але при цьому зрив спостереження або неоднозначність вимірювань не спостерігається.

Роботу шестиантенного пеленгатору краще пояснити на прикладі його включення в систему автоматичного супроводження за кутовими координатами. Така система починає працювати після вирішення задачі пошуку кутового положення джерела радіовипромінювання. Тобто на начальному етапі ціль вже знаходиться у межах робочого діапазону кутів радіопеленгатору. Наприклад кутове положення цілі дорівнює 110°, тоді на виході пеленгатору буде позитивне значення напруги, що в подальшому приведе в дію антено-поворотний механізм для суміщення кутів $\theta_b$ і $\theta_s$. З кожним новим вимірюванням напруга буде спадати до моменту поки кути не будуть дорівнювати і на виході матимемо значення напруги 0. Суцільна чорна крива на рис. 6 від значення кута 110° до 90° маємо ділянки з різною крутизно. Найвища крутизна навколо рівносигнального напрямку дасть найвищу точність вимірювань, інші ділянки матимуть меншу точність. В той самий час від 110° до 90° маємо спадну криву без множинних максимумів, на відміну від пеленгаційної кривої двоантенного пеленгатору. Цей факт показує, що з кожним новим циклом вимірювань система автоматичного супроводження за кутовими координатами прийде до істинного напрямку на $\theta_s$. Перевага шестиантенного над чотирьохантенним проявляється саме на кінцевому етапі слідкування за $\theta_s$. При цьому двоантенний має найвищу крутизну, але у дуже малому діапазоні кутів. Можна стверджувати, що шестиантенний пеленгатор має компромісні характеристики між точністю та діапазоном кутів однозначних вимірювань і дозволяє подолати існуюче протиріччя.

## 5 РЕЗУЛЬТАТИ

Перевірити отримані результати також доцільно на прикладі аналізу виразів для граничних похибок вимірювання радіопеленгаторів. В статистичній теорії


DOI 10.15588/1607-3274-2024-4-3

оптимізації радіотехнічних систем [16] граничну дисперсію похибки оцінювання параметра $\lambda$ знаходять з нерівності Крамера-Рао

$$\sigma_{\lambda}^2 \geq -\frac{1}{< \frac{d^2}{d\lambda^2} \ln P[\bar{u}(t) \mid \lambda] >} \Bigg|_{\lambda=\lambda_{true}}, \qquad (29)$$

де $< \cdot >$ – знак статистичного усереднення, $d^2 / d\lambda^2$ – символ вторинної похідної.

Загальний аналітичний вираз для граничної похибки оцінки $\theta_s$ всіх радіопеленгаторів має наступний вигляд:

$$\sigma_{\theta_s}^2 = \left( \frac{2}{N_0} \sum_{i=1}^{N} \int_T \left( \mathrm{Re} \left\{ \frac{d\dot{s}_i(t,\theta_s)}{d\theta_s} \right\} \right)^2 dt \right)^{-1} \Bigg|_{\theta_s=\theta_{s\ true}} \cdot \qquad (30)$$

Граничні дисперсії похибок оцінювання кутового положення джерела для розглянутих в цій статті випадків мають наступний вигляд:

1) для двохантенного амплітудного пеленгатора

$$\sigma_{\theta_s}^2 = \frac{1}{\mu_0} \left\{ \left( \frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 \right\}^{-1} \Bigg|_{\theta_s=\theta_{s\ true}},$$

2) для чотирьохантенного амплітудного пеленгатора

$$\sigma_{\theta_s}^2 = \frac{1}{\mu_0} \left\{ \left( \frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_2(\theta-\theta_b-0,5\theta_{\delta 2})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_2(\theta-\theta_b+0,5\theta_{\delta 2})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 \right\}^{-1} \Bigg|_{\theta_s=\theta_{s\ true}},$$

3) для шестиантенного амплітудно-фазового пеленгатора

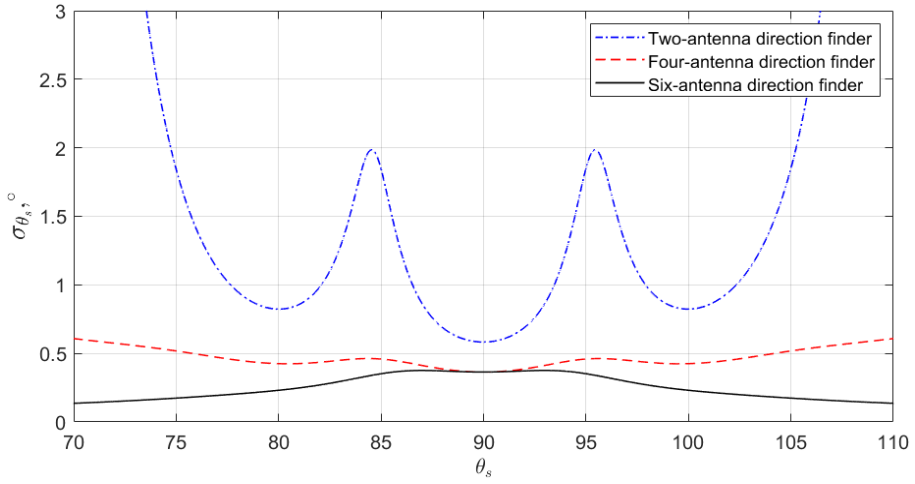
$$\sigma_{\theta_s}^2 = \frac{1}{\mu_0} \left\{ \left( \frac{dG_1(\theta-\theta_b-0,5\theta_{\delta 1})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_1(\theta-\theta_b+0,5\theta_{\delta 1})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_2(\theta-\theta_b-0,5\theta_{\delta 2})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( \frac{dG_2(\theta-\theta_b+0,5\theta_{\delta 2})}{d\theta} \Bigg|_{\theta=\theta_s} \right)^2 + \left( K_{05}^2 + K_{06}^2 \right) \times \left( \pi f_0 dc^{-1} \right)^2 G_3^2(\theta_0) \sin^2(\theta_s-\theta_b) \right\}^{-1} \Bigg|_{\theta_s=\theta_{s\ true}},$$

де $\mu_0 = 0,5 E_s N_0^{-1}$.

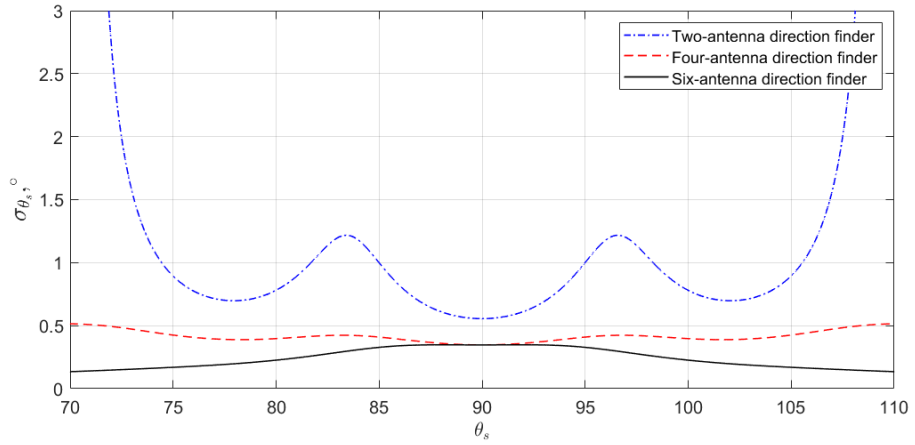
Отримані вирази доцільно промоделювати при тих самих параметрах, що і пеленгаційні характеристики на рис. 6. Результат такого моделювання наведено на рис. 7. На рисунках зображень середньоквадратичні похибки оцінювання кутового положення джерела радіовипромінювання у градусах в залежності від кута спостереження для двох типів діаграм (25) і (26). Обрані кути від 70° до 110° відповідають найширшому діапазону однозначних вимірювань в шестиантенному пеленгаторі.

З аналізу отриманих залежностей слідує, що в двохантенному радіопеленгаторі найвища точність зосереджена на лінійних ділянках пеленгаційних характеристик. Чотирьохантенний пеленгатор також має коливання граничних похибок, але при цьому за рахунок збільшеної кількості незалежних вимірювань ці коливання незначні і загалом граничні похибки менші ніж в двохантенному. Шестиантенний пеленгатор має найвищу точність, що за рахунок доданку у вигляді швидкоспадної функції $\sin^2(\theta_s - \theta_b)$ значно спадає за межами рівносигнального напрямку. В напрямку $\theta_b$, тільки в одній точці, граничні похибки шестиантенного пеленгатору співпадають з чотирьох антенним, адже $\sin^2(0) = 0$.


DOI 10.15588/1607-3274-2024-4-3




a




b

Рисунок 7 – Середньоквадратичні похибки оцінювання кутового положення джерела радіовипромінювання (блакитний – двоантенний, червоний – чотирьоантенний, чорний – шести антенний пеленгатор): а – діаграми спрямованості у вигляді функцій Гауса, б – діаграми спрямованості у вигляді Sinc-функцій

## ВИСНОВКИ

**Наукова новизна** цього дослідження полягає в розробці та синтезі нового методу обробки сигналів, що забезпечує високоточне та однозначне визначення кутових положень джерел радіовипромінювання в багатоантенних радіопеленгаторах.

**Практична значимість** полягає у можливості застосування розроблених методів для підвищення ефективності роботи бортових систем безпілотних літальних апаратів.

**Для майбутніх досліджень** важливо врахувати, що отримані алгоритми працюють за умови повної відомості моделі сигналу. Тому перспективним напрямком є дослідження методів пеленгації стохастичних сигналів, які характеризуються лише модулем та кореляційними характеристиками. Це дозволить розширити можливості застосування розроблених методів у більш складних умовах роботи радіопеленгаторів.

## ПОДЯКИ

Робота виконана за фінансової підтримки Національного фонду досліджень України, за сприяння Кембриджського університету (Велика Британія), державний реєстраційний номер проекту 0124U003671.

## ЛІТЕРАТУРА

1. Improving Surface Current Resolution Using Direction Finding Algorithms for Multiantenna High-Frequency Radars/ [B. An, J. Yim, S. Shin, C. Kim, D. Kim] // IEEE Transactions on Geoscience and Remote Sensing. – 2020. – Vol. 58, № 7. – P. 4951–4960. DOI: 10.1109/TGRS.2020.2967439.
2. Interoperability of Direction-Finding and Beam-Forming High-Frequency Radar Systems: An Example from the Australian High-Frequency Ocean Radar Network / [M. L. Heron, A. Prytz, R. Bailey, T. Abhayapala] // Remote Sensing. – 2021. – Vol. 13, № 9. – P. 1814. DOI: 10.3390/rs13091814.
3. Wyatt L.R. Marine Applications of High-Frequency Direction-Finding Radar: Detection and Monitoring of Surface Targets/ L. R. Wyatt, C. J. Lwyarczyk, P. Robinson // Marine Technology Society Journal. – 2020. – Vol. 54, № 5. – P. 23–33. DOI: 10.4031/MTSJ.54.5.2.
4. Stankovic S. Advances in Direction Finding Algorithms for Electronic Warfare/ S. Stankovic, I. Djurovic // IEEE Aerospace and Electronic Systems Magazine. – 2019. – Vol. 34, № 8. – P. 58–72. DOI: 10.1109/MAES.2019.295345.
5. Wang L. Advanced Direction Finding Systems: Frequency Range, Accuracy, and Measurement / L. Wang, M. Zhao //


DOI 10.15588/1607-3274-2024-4-3

Electronics Research Letters. – 2021. – Vol. 18, № 2. – P. 112–126. DOI: 10.3390/erl18020112.
6. Lee C. Direction-Finding Techniques and System Performance / C. Lee, S. Kim // International Journal of Electromagnetic and Microwave Studies. – 2019. – Vol. 15, № 4. – P. 78–92. DOI: 10.3390/ijems1504078.
7. Smith J. Trends and Applications of Unmanned Aerial Vehicles: A Review/ J. Smith, A. Brown // Journal of Unmanned Systems. – 2023. – Vol. 12, № 4. – P. 234–250. DOI: 10.3390/jus12040234.
8. Lee K. Advancements in Navigation and Positioning Systems for Autonomous Vehicles / K. Lee, Y. Zhang // Sensors. – 2024. – Vol. 24, № 6. – P. 1520–1535. DOI: 10.3390/s24061520.
9. Johnson M. Improving Angular Position Determination in Radio Systems for Autonomous Platforms/ M. Johnson, S. Patel // IEEE Transactions on Aerospace and Electronic Systems. – 2024. – Vol. 60, № 2. – P. 674–689. DOI: 10.1109/TAES.2024.1234567.
10. Williams D. Amplitude Comparison Direction Finding Systems for Modern Radar Applications / D. Williams, P. O'Connell // International Journal of Microwave and Wireless Technologies. – 2021. – Vol. 13. – P. 125–138. DOI: 10.1017/S1759078720001530.
11. Kumar V. Interferometric Techniques in Modern Radar Systems / V. Kumar, R. Patel // IEEE Transactions on Antennas and Propagation. – 2022. – Vol. 70. – P. 985–999. DOI: 10.1109/TAP.2022.3141859.
12. Garcia M. Correlation-Interferometric Direction Finding in Complex Environments / M. Garcia, J. Lee // Electronics. – 2023. – Vol. 12. – P. 1123. DOI: 10.3390/electronics12051123.
13. Volosyuk V. Statistical Theory of Optimal Functionally Deterministic Signals Processing in Multichannel Aerospace Imaging Radar Systems/ V. Volosyuk, S. Zhyla // Computation. – 2022. – Vol. 10. – P. 213. DOI: 10.3390/computation10120213.
14. Volosyuk V. Statistical Theory of Optimal Stochastic Signals Processing in Multichannel Aerospace Imaging Radar Systems / V. Volosyuk, S. Zhyla // Computation. – 2022. – Vol. 10. – P. 224. DOI: 10.3390/computation10120224.
15. Волосюк В.К. Статистическая теория радиотехнических систем дистанционного зондирования и радиолокации / В. К. Волосюк, В. Ф. Кравченко. – М. : Физматлит, 2008. – 704 с.
16. Van Trees H.L. Detection, Estimation, and Modulation Theory, Part I. / H. L. Van Trees. – New York : John Wiley & Sons, 2001. – 716 p.

Стаття надійшла до редакції 20.09.2024.
Після доробки 10.11.2024.

UDC 621.396

## OPTIMIZATION OF RADIO SYSTEM STRUCTURES FOR DIRECTION FINDING OF SIGNAL SOURCES WITH COMPLETELY KNOWN PARAMETERS

**Zhyla S. S.** – Dr. Sc, Associate Professor, Head of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

**Tserne E. O.** – PhD, Junior Researcher of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

**Popov A. V.** – Dr. Sc, Associate Professor, Professor of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

**Ruzhentsev N. V.** – Dr. Sc, Professor, Chief Scientist of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

**Volkov Ye. G.** – Director, State Enterprise "Research Institute "Orion", Kyiv, Ukraine.

**Shevchuk S. D.** – Deputy Chief of Department No. 91, State Enterprise "Research Institute "Orion", Kyiv, Ukraine.

**Gribsky O. P.** – Senior Researcher, State Enterprise "Research Institute "Orion", Kyiv, Ukraine.

**Kolesnikov D. V.** – PhD, Senior Lecturer of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

**Inkarbaieva O. S.** – PhD, Senior Lecturer of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

**Cherepnin G. S.** – PhD, Senior Lecturer of the Department of Aerospace Radio-Electronic Systems, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

### ABSTRACT

**Context.** Direction finders are critical components of radar and radio navigation systems, particularly when installed onboard UAVs. The increasing use of unmanned systems has heightened the need for precise direction finding and wide-angle unambiguous measurements. The primary challenge is to strike a balance between achieving high accuracy and maintaining a broad range of unambiguous measurement angles.

**Objective.** To simultaneously enhance direction finding accuracy and expand the range of unambiguous measurement angles through the statistical synthesis of functionally deterministic signal processing methods in multichannel direction finders.

**Method.** The approach is grounded in the statistical theory of optimization for radio remote sensing and radar systems. For the specific type of signals considered in this study, represented by functional-deterministic models, the likelihood function is constructed, and its maxima are determined for various configurations of multi-antenna direction finders. The statistical synthesis results are validated through simulation and in-situ experiments.

**Results.** Theoretical analysis and simulation modeling confirm that in dual-antenna direction finders, there is a trade-off between high resolution and the range of unambiguous direction finding angles. An improved signal processing method is developed for a four-antenna direction finder, utilizing a pair of high-gain and a pair of low-gain antennas. To achieve the maximum possible bearing accuracy within the range of unambiguous direction finder measurements, a new signal processing method is synthesized for a six-element radio receiver, combining the processing of signals in two amplitude direction finders and one phase direction finder.

**Conclusions.** The proposed approach achieves an optimal balance between resolution and angle range, making it particularly suitable for onboard systems of unmanned aerial vehicles. Simulation results confirm the effectiveness of the proposed method, highlighting its potential for implementation in modern radio systems.

**KEYWORDS:** multi-antenna direction finders, statistical optimization, optimal signal processing, simulation modeling.


DOI 10.15588/1607-3274-2024-4-3

**REFERENCES**

1. An B., Yim J., Shin S., Kim C., Kim D. Improving surface current resolution using direction finding algorithms for multiantenna high-frequency radars, *IEEE Transactions on Geoscience and Remote Sensing*, 2020, Vol. 58(7), pp. 4951–4960. DOI: 10.1109/TGRS.2020.2967439.
2. Heron M.L., Prytz A., Bailey R., Abhayapala T. Interoperability of direction-finding and beam-forming high-frequency radar systems: An example from the Australian High-Frequency Ocean Radar Network, *Remote Sensing*, 2021, Vol. 13(9), P. 1814. DOI: 10.3390/rs13091814.
3. Wyatt L.R., Lwyarczyk C.J., Robinson P. Marine applications of high-frequency direction-finding radar: Detection and monitoring of surface targets, *Marine Technology Society Journal*, 2020, Vol. 54(5), pp. 23–33. DOI: 10.4031/MTSJ.54.5.2.
4. Stankovic S., Djurovic I. Advances in direction finding algorithms for electronic warfare, *IEEE Aerospace and Electronic Systems Magazine*, 2019, Vol. 34(8), pp. 58–72. DOI: 10.1109/MAES.2019.295345.
5. Wang L., Zhao M. Advanced direction finding systems: Frequency range, accuracy, and measurement, *Electronics Research Letters*, 2021, Vol. 18(2), pp. 112–126. DOI: 10.3390/erl18020112.
6. Lee C., Kim S. Direction-finding techniques and system performance, *International Journal of Electromagnetic and Microwave Studies*, 2019, Vol. 15(4), pp. 78–92. DOI: 10.3390/ijems1504078.
7. Smith J., Brown A. Trends and applications of unmanned aerial vehicles: A review, *Journal of Unmanned Systems*, 2023, Vol. 12(4), pp. 234–250. DOI: 10.3390/jus12040234.
8. Lee K., Zhang Y. Advancements in navigation and positioning systems for autonomous vehicles, *Sensors*, 2024, Vol. 24(6), pp. 1520–1535. DOI: 10.3390/s24061520.
9. Johnson M., Patel S. Improving angular position determination in radio systems for autonomous platforms, *IEEE Transactions on Aerospace and Electronic Systems*, 2024, Vol. 60(2), pp. 674–689. DOI: 10.1109/TAES.2024.1234567.
10. Williams D., O'Connell P. Amplitude comparison direction finding systems for modern radar applications, *International Journal of Microwave and Wireless Technologies*, 2021, Vol. 13, pp. 125–138. DOI: 10.1017/S1759078720001530.
11. Kumar V., Patel R. Interferometric techniques in modern radar systems, *IEEE Transactions on Antennas and Propagation*, 2022, Vol. 70, pp. 985–999. DOI: 10.1109/TAP.2022.3141859.
12. Garcia M., Lee J. Correlation-interferometric direction finding in complex environments, *Electronics*, 2023, Vol. 12, 1123. DOI: 10.3390/electronics12051123.
13. Volosyuk V., Zhyla S. Statistical theory of optimal functionally deterministic signals processing in multichannel aerospace imaging radar systems, *Computation*, 2022, Vol. 10, 213. DOI: 10.3390/computation10120213.
14. Volosyuk V., Zhyla S. Statistical theory of optimal stochastic signals processing in multichannel aerospace imaging radar systems, *Computation*, 2022, Vol. 10, 224. DOI: 10.3390/computation10120224.
15. Volosyuk V.K., Kravchenko V.F. Statisticheskaya teoriya radiotekhnicheskikh sistem distantsionnogo zondirovaniya i radiolokatsii, *Fiziko-Matematicheskaya Literatura*, Moscow, Russia, 2008, 704 p.
16. Van Trees H.L. Detection, estimation, and modulation theory, Part I, *John Wiley & Sons*, New York, NY, USA, 2001.


DOI 10.15588/1607-3274-2024-4-3

# МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

# MATHEMATICAL AND COMPUTER MODELING

UDC 004.94

## MODELING OF THE SPREAD OF TUBERCULOSIS BY REGIONS IN UKRAINE

**Boyko N. I.** – PhD, Associate Professor, Associate Professor of the Department of Artificial Intelligence Systems, Lviv Polytechnic National University, Lviv, Ukraine.

**Rabotiahov D. S.** – Student, Department of Artificial Intelligence Systems, Lviv Polytechnic National University, Lviv, Ukraine.

### ABSTRACT

**Context.** Modelling the spread of tuberculosis in Ukraine is particularly relevant due to the increasing number of cases, especially in 2023.

**Objective.** The aim of this study is to solve modeling tasks by applying modern machine learning methods and data analysis to build predictive models of tuberculosis spread at the regional level.

**Method.** To model the spread of tuberculosis at the regional level in Ukraine, it is proposed to use several approaches, such as the SIR model, cellular automata, and Random Forest. Each of these methods has its unique advantages and can provide a more detailed understanding of the dynamics of disease spread. The SIR model (Susceptible-Infectious-Recovered) is a classical epidemiological model that describes the spread of infectious diseases in a population. The model assumes three groups of the population: *S* (Susceptible) – susceptible to infection; *I* (Infectious) – infected and capable of transmitting the infection; *R* (Recovered) – those who have recovered and gained immunity. Cellular automata are a discrete model that uses a grid of cells to simulate spatiotemporal processes. Each cell can be in different states (e.g., healthy, infected, immune) and change its state depending on the states of neighboring cells. Random Forest is a machine learning method that uses an ensemble of decision trees for classification or regression. This method can be applied to predict the spread of tuberculosis based on a large number of input parameters. Using these methods will allow for a deep analysis and comprehensive results regarding the spread of tuberculosis at the regional level in Ukraine. This, in turn, will facilitate the development of effective strategies to combat the disease and improve public health..

**Results.** The results of applying the Random Forest and SIR methods were described and analyzed in detail. For Random Forest, the metrics MSE and $R^2$ were evaluated, showing high prediction accuracy. In the case of the SIR algorithm, visual assessment of the results revealed insufficient accuracy due to model limitations. Comparing the chosen methods with other studies, a conclusion was made about the need to consider more complex algorithms to obtain more accurate results.

**Conclusions.** Based on the research results, it can be concluded that the Random Forest method is sufficiently effective for predicting vulnerable social groups and that the SIR algorithm is less effective for modeling the spread of tuberculosis. For further research development, it is recommended to consider more complex algorithms and account for additional factors influencing the spread of the disease. Moreover, to better understand further actions to combat the disease, it is advisable to simulate the spread of tuberculosis among the population of Ukraine.

**KEYWORDS:** method, metric, tuberculosis, Random Forest, Susceptible-Infectious-Recovered, modeling, algorithm.

### ABBREVIATIONS

TB is a tuberculosis;
WHO is a World Health Organization;
DT is a Decision Tree;
SIR is a Susceptible-Infectious-Recovered;
MSE is a Mean Squared Error.

### INTRODUCTION

Modeling the spread of tuberculosis in Ukraine at the regional level in 2024 is particularly relevant due to the increasing number of cases, especially in 2023. Tuberculosis remains one of the significant global problems of modern society, particularly in the context of the pandemic, and requires a comprehensive approach to its study and control [1, 3].

The scientific development of this problem includes various methods and approaches. Currently, statistical methods, epidemiological models, and machine learning methods are already being used to analyze and predict the spread of the disease. However, there are significant gaps in understanding the dynamics and specifics of tuberculosis spread at the regional level in Ukraine [5].

**Object of the study:** The social and demographic aspects of tuberculosis spread by gender and age.

**Subject of the study:** The methods and algorithms for studying the spread of tuberculosis in the regions of Ukraine.


DOI 10.15588/1607-3274-2024-4-4

**The aim of this study** is to address these issues by applying modern machine learning methods and data analysis to build predictive models of tuberculosis spread at the regional level. The conclusions drawn from this research can make a significant contribution to the development of effective strategies for disease control and prevention in Ukraine.

**Tasks of the research:**

– Data collection and preparation: Assess the available data on the number of tuberculosis cases in each region of Ukraine for 2024, considering age and gender distribution.

– Data analysis: Study the distribution of tuberculosis cases by region, age, and gender to identify possible dependencies and correlations.

– Develop a predictive model: Apply machine learning methods, particularly the Random Forest algorithm, to build a predictive model for tuberculosis spread at the regional level. Consider risk factors such as age and gender to identify the most vulnerable population groups.

– Evaluate model effectiveness: Analyze the results and evaluate the accuracy and reliability of the predictive model. Identify the most vulnerable population groups.

The advantages of modeling for such tasks lie in the ability to predict disease spread dynamics, identify the most at-risk population groups, and determine effective control strategies. Models allow for the consideration of various factors, such as demographic and socio-economic characteristics, which helps in understanding the complex relationships affecting disease spread. They can also be a useful tool for decision-making and developing preventive and treatment strategies.

The relevance of the topic reflects the importance and significance of the problem being studied and its alignment with contemporary scientific and practical needs. Modeling the spread of tuberculosis in Ukraine is particularly relevant due to the increasing number of cases, especially in 2023. The study is driven by the high incidence rate and potential threat to public health, which requires thorough analysis and effective management strategies [2].

Understanding the spread of tuberculosis and identifying the most vulnerable population groups is crucial for further control and prevention of this disease. Research in this area not only provides a scientific component but also has a direct practical impact on the health of citizens and the healthcare system of the country [6].

Moreover, modeling the spread of tuberculosis with the identification of the most vulnerable population groups by gender and age is significant for the further development of medical science and practice. Discovering new connections and factors influencing disease spread can contribute to improving tuberculosis diagnosis and treatment methods, as well as developing effective control and prevention programs.

The main focus of the research is on identifying the connections between various social and demographic factors and the spread of the disease, as well as determining the factors contributing to the risk of illness among different population groups. Thus, the object and subject of the research reflect the key aspects investigated within this work to address the problem of tuberculosis spread and improve public health.

## 1 PROBLEM STATEMENT

The relevance of this research lies in its ability to forecast disease spread dynamics, identify the most at-risk population groups, and determine effective control strategies. The models developed will consider various factors, such as demographic and socio-economic characteristics, to understand the complex relationships affecting disease spread. These models can also serve as a valuable tool for decision-making and developing preventive and treatment strategies.

Let:

– $N$ be the total population of a region;

– $S(t)$ be the number of susceptible individuals at time $t$;

– $I(t)$ be the number of infectious individuals at time $t$;

– $R(t)$ be the number of recovered individuals at time $t$;

– $\beta$ be the transmission rate;

– $\gamma$ be the recovery rate.

The **SIR** model is described by the following set of differential equations [5, 7]:

$$\frac{dS(t)}{dt} = -\beta \frac{S(t)I(t)}{N},$$

$$\frac{dI(t)}{dt} = \beta \frac{S(t)I(t)}{N} - \gamma I(t),$$

$$\frac{dR(t)}{dt} = \gamma I(t).$$

Objective Function for SIR Model: Minimize the error between the model predictions and the actual data [8, 9]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[I_{actual}(t_i) - I_{predicted}(t_i)]^2.$$

For the **Random Forest** model, let:

– $X$ be the feature matrix (including age, gender, socio-economic factors, etc.),

– $y$ be the target variable (number of tuberculosis cases).

The objective is to train the Random Forest model to minimize the mean squared error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[y_i - \hat{y}_i]^2,$$

where $\hat{y}_i$ is the predicted value from the Random Forest model.


DOI 10.15588/1607-3274-2024-4-4

Using these methodologies, we aim to develop accurate predictive models to understand the spread of tuberculosis and identify the most vulnerable groups, thereby contributing to better disease control and prevention strategies.

## 2 LITERATURE REVIEW

To perform this research there will be used various methods aimed at analysing the socio-demographic aspects of the spread of tuberculosis and identifying the most vulnerable groups of the population. One of the main methods is the analysis of statistical data, which will provide objective results on the distribution of the disease among different categories of the population. Machine learning methods such as regression or clustering algorithms will also be used to identify complex relationships and patterns in the data. To analyse the geographical distribution of the disease, datasets on different regions over a long period of time will be used. In addition, epidemiological models will be used to predict the spread of tuberculosis and evaluate the effectiveness of control strategies. Such a comprehensive approach to research will allow us to gain a deeper understanding of the problem and identify the best ways to combat the disease.

Ilnytskyi H. I. (2021) [4, 10] in his research work investigated the epidemiological situation with tuberculosis in Ukraine. In addition to analysing statistical data and conducting sociological research, the author also developed a mathematical cellular SIS model to predict the spread of tuberculosis in Ukraine. The methods are based on a system of differential equations that describe the dynamics of tuberculosis spread in a population. The methods take into account factors such as TB incidence, TB mortality, infection rates and treatment effectiveness. This approach is not very suitable for the real world, as it assumes that an individual becomes infectious immediately after infection, which is not true, as the disease has an incubation period. The author also considered the SEIS model. The SEIS model considered here can be interpreted as an SIS model with an effective delay in the spread of the disease. It performed better under conditions close to those of the real world.

Researcher Rogovskyi V. O. (2023) [11, 18] developed a mathematical model for predicting the success of tuberculosis treatment. The model is based on a system of differential equations that describe the dynamics of the number of tuberculosis pathogens in the body. The model takes into account the following factors: the sensitivity of the pathogen to chemotherapeutic drugs, the patient's immune status, and the patient's compliance with the chemotherapy regimen. As analysed by the author, the model makes good predictions for short periods of time, as evidenced by the margin of error, but applying this model over longer periods of time can lead to an increase in the prediction error.

According to the Public Health Centre of the Ministry of Health of Ukraine [5, 12], 23,788 new cases of tuberculosis were registered in 2022, which is 60.1 per 100,000 people. Expectations for reducing TB morbidity and mortality in 2022: 10% reduction in morbidity and 5% reduction in mortality. Actual results: 8% reduction in incidence and 7% reduction in mortality.

The section "Creating a Bayesian Network of Risk Factors for COVID-19" of the thesis by Yaroslav Shevchenko [13, 15] describes two methods used to build a Bayesian network: the method of expert opinion (the author interviews 10 experts (infectious disease doctors) about their opinion on the impact of 12 risk factors on the likelihood of contracting COVID-19). Based on these opinions, a Bayesian network is built. Pros: simplicity and accessibility (the expert opinion method is easy to use and does not require special programming knowledge), flexibility (Bayesian networks can easily adapt to new data and information), visuality (Bayesian networks provide a convenient visual representation of the relationships between risk factors). Cons: subjectivity (the method of expert assessment can be subjective, depending on the opinion of experts), data requirements (the training algorithm of the Bayesian network requires a large amount of data), difficulty of interpretation (interpretation of the results of the Bayesian network can be difficult for people without special knowledge).

From the journal Chemistry, Ecology and Education, section "Mathematical modelling of the spread of viral infections in local urban ecosystems" [17, 20], one can learn about the process of spreading respiratory viral diseases, which are most often transmitted by airborne droplets. Infection by this route is most likely in crowded places. It is important to note that people start infecting others before they become ill. The following modelling techniques were used: a mathematical model (the authors developed a mathematical model to predict the spread of respiratory viral infections), Langevin dynamics (the model uses Langevin dynamics to model the movement of agents in the system), Levy flight modelling (the model allows for the sudden movement of infected agents over long distances). Advantages: adequacy (the model adequately reflects the main spatial and temporal components of the urban ecosystem), versatility (the model is universal and can be adjusted according to needs), effectiveness (the model can be used to predict epidemics and evaluate the effectiveness of prevention methods). Cons: Complexity (the model can be difficult to understand and use), need for the large amounts of data (data on the urban ecosystem and human behaviour are required to parameterise the model), limitations (the model cannot take into account all factors that affect the spread of infections).

Artificial intelligence can be a powerful tool for predicting the spread of TB and developing effective control strategies. An article titled "Prospects for the application of artificial intelligence to predict the spread of tuberculosis infection in the WHO European Region" [14, 19] only confirms this. In the study, the authors describe various AI methods that can be used to predict TB. TB spread models, that were mentioned in the paper:

– the classic SIR model:

– The model uses three states for agents: susceptible, infected, and recovered.


DOI 10.15588/1607-3274-2024-4-4

– The model takes into account factors such as the number of agents, speed of movement, probability of infection, duration of the disease, etc.

– Urban environment model:

– The authors propose to develop a model that takes into account the life, behaviour and interaction of people in the city.

– This model will be integrated with the infection spread model for more accurate forecasting.

– Advantages of the proposed approach:

– Better forecasting: The combination of models will allow for more accurate forecasting of the spread of the epidemic at the regional, national and global levels.

– Incorporating geospatial data: The use of geographic maps and the location of buildings will allow us to study the spatial spread of the epidemic.

– Speed of calculations: The model should be fast enough to run a large number of computer experiments.

– Parallel computing: The possibility of parallel computing will allow modelling the spread of the epidemic at the macro level of countries and the world.

– Disadvantages of the proposed approach:

– Complexity of development: Developing an urban environment model is a complex task.

– The model must take into account many factors such as age, immunity, building types, etc.

– Data requirements: Large amounts of data are required to develop and train the model.

Table 1 provides an analysis of the research on this topic, which is subject to careful scientific review in order to understand the direction of change.

Table 1 – Review of related papers

| Title of the work (author) | Methodology | Pros of the methodology | Cons of the methodology |
|---|---|---|---|
| Ilnytskyi H. I. (2021) | Mathematical cellular model SIS/SEIS | Simple, describes the dynamics of TB spread well | Does not take into account the incubation period, not very suitable for the real world |
| Rogovskyi V. O. (2023) | Mathematical model of the dynamics of the number of TB pathogens | Allows to take into account susceptibility to chemotherapy, immune status, adherence to the regimen | Not very accurate over long periods of time |
| Public Health Centre of the Ministry of Health of Ukraine (2022) | Expectations and actual results in reducing morbidity and mortality from tuberculosis | A simple method of forecasting | Does not take into account the impact of various factors |
| Shevchenko Yaroslava (2020) | Bayesian network of risk factors for COVID-19 | Simple, flexible, visual | Subjective, requires a lot of data, difficult to interpret |

Summing up the results presented in Table 1, the analysis of the literature confirms the relevance of the problem of studying tuberculosis in Ukraine. Despite a certain decline in the incidence rate in recent years, the country's TB rate remains higher than in the European Union.

## 3 MATERIALS AND METHODS

In order to predict the spread of tuberculosis and to model it, we need a dataset that contains sufficient information to make the predictions accurate. That is why the dataset [17, 20] was chosen for the paper, which contains data on the incidence of tuberculosis in Ukraine for a long period from 2007–2022. The data is presented in tabular format and describes the following characteristics: the number of cases by date and region, age and gender, TB form and treatment outcome.

Although the dataset does not provide direct information on the causes of TB, it does allow for the study of risk factors such as age, place of residence, socioeconomic status, and comorbidities.

The data describe both men and women of all age groups, with detailed age categorisation.

In addition to information on the forms of TB, the dataset also contains data on treatment outcomes, which allows for an assessment of treatment effectiveness.

The existing dataset will be split 7:3 for training and testing, respectively.

For a better understanding of the approaches and to solve the tasks, let us consider the proposed methods listed in Table 1

The methodology of the mathematical cellular model SIS/SEIS described by Ilnytskyi G. I. (2021) has the advantages of ease of use and a good description of the dynamics of the disease spread, however, an important negative factor of the methodology is that it does not take into account the incubation period of the disease, which is why its use in this work is inappropriate [16].

Researcher Rogovsky V. O. (2023) proposed a mathematical model of the dynamics of the number of tuberculosis pathogens, which perfectly allows taking into account sensitivity to chemotherapy, immune status, and compliance with the regimen, but the lack of need to include these parameters in the analysis, along with the insufficient accuracy of predictions over long periods of time, make this methodology inappropriate for use in this work.

The data provided by the Public Health Centre of the Ministry of Health of Ukraine (2022) only shows comparative statistics on the expected and actual results of reducing the incidence and mortality from tuberculosis.

To address this issue, Yaroslav Shevchenko's thesis proposes a Bayesian network of risk factors for contracting COVID-19. The advantages of this method are the flexibility of the model and a good opportunity to visualise the results, but the model requires a large data set, which contradicts the above description of the data set, and the results of the algorithm are difficult to interpret.


DOI 10.15588/1607-3274-2024-4-4

The problem of the need for a large data set was also encountered by the methods of Langevin dynamics and Levy flight modelling described in the journal Chemistry, Ecology and Education (2023), although they had such advantages as the versatility of the algorithm and its efficiency.

The article “Prospects for the application of artificial intelligence to predict the spread of tuberculosis infection in the WHO European Region” presents the classic SIR model, an urban environment model that has such advantages as the inclusion of geospatial data, calculation speed, and the possibility of parallel computing. There are also disadvantages of the methodology, including the complexity of development and the need for a medium to large data set.

However, for a comparative analysis, the methods described in Table 1 should be considered, as well as additional tools that are best suited to the task at hand.

To address the issue of predicting the most vulnerable populations by age and gender in terms of TB incidence, decision trees can be used, namely Random Forest. Random Forest is a machine learning algorithm that uses a combination of DTs to improve accuracy. It works well with medium-sized datasets, because in our case, the selected dataset contains information for each of the 24 regions of Ukraine and the city of Kyiv in the time period from 2007–2022, which makes it possible to track the dynamics of the disease in each of the regions over an average period of 15 years.

Random Forest can help us to:

– Identify risk factors: Random Forest can help identify factors (gender, age, other) that influence the likelihood of getting TB.

– Classifying people into risk groups: Random Forest can classify people into risk groups based on their likelihood of getting the disease.

– Identify the most vulnerable groups by gender and age: Random Forest can help you identify the populations that are most at risk for TB.

The model will make predictions about the most vulnerable groups based on the number of cases, age groups, gender, and year. As a result of processing the data, the model will output the age group and gender of people who may be most affected by the disease.

The model works as follows:

– Random sampling: A set of random subsets is generated from the data.

– Training of the DTs: A DT is trained for each subset of the data.

– Aggregation: The forecasts from all the DTs are combined to produce the final forecast.

Here is the pseudo-code of the algorithm:

```
Algorithm Random Forest: pseudocode
1:  To generate c classifiers:
2:  for i = 1 to c do
3:  Randomly sample the training data D with
    replacement to produce Di;
4:  Create a root node, Ni, containing Di,
5:  Call BuildTree(Ni)
6:  end for
7:  BuildTree(N):
8:  if N contains instances of only one class
    then
9:  return
10: else
11: Randomly select x% of the possible split-
    ting features in N
12: Select the feature F with the highest in-
    formation gain to split on
13: Create f child nodes of N, N1,.... , Nf,
    where F has f possible values (F1, ...,
    Ff)
14: for i = 1 to f do
15:      Set the contents of Ni, to Di, where
    Di, is all instances in N that match Fi
16:   Call BuildTree(Ni)
17: end for
18: end if
```

Random Forest algorithm:

– You need to set the following parameters:

– Number of trees;

– Depth of trees.

– Next, you need to train the model, for this(for each tree):

– Select a random subset of the data;

– Train the DT on the subset of data.

– The next step is to make a prediction, for this(for each tree):

– Make a prediction for a new data instance;

– Combine the predictions from all the trees.

Also, for a better understanding of how the algorithm works, we can look at the flowchart of its operation in Fig. 1.

In Fig. 1 the following notation is used:

$V_n$: This is the entire dataset used to train the algorithm. It consists of $n$ data points, where each data point is represented by a pair ($x_i$, $y_i$). Here, xi is the feature vector for the ith data point and $y_i$ is the corresponding objective value.

$v$: This is the number of decision trees to be created in the algorithm.

$V_1$, $V_2$, ..., $V_k$: These are the random data sets used to train each decision tree. Each data set $V_i$ consists of $k$ data points that are randomly selected from $V_n$.

$k$: This is the number of features that are randomly selected for each decision tree.

$A_1$, $A_2$, ..., $A_k$: These are the predictions made by each decision tree for a new data instance $X$.

$A$: This is the final prediction made by the Random Forest algorithm. It is calculated as the average of the predictions made by all decision trees.

The $R^2$-measure (coefficient of determination) can be used to assess the accuracy, and the mean square error (the average of the squared differences between the predictions and the actual values), which will be referred to as MSE, can be used to assess the model error.


DOI 10.15588/1607-3274-2024-4-4

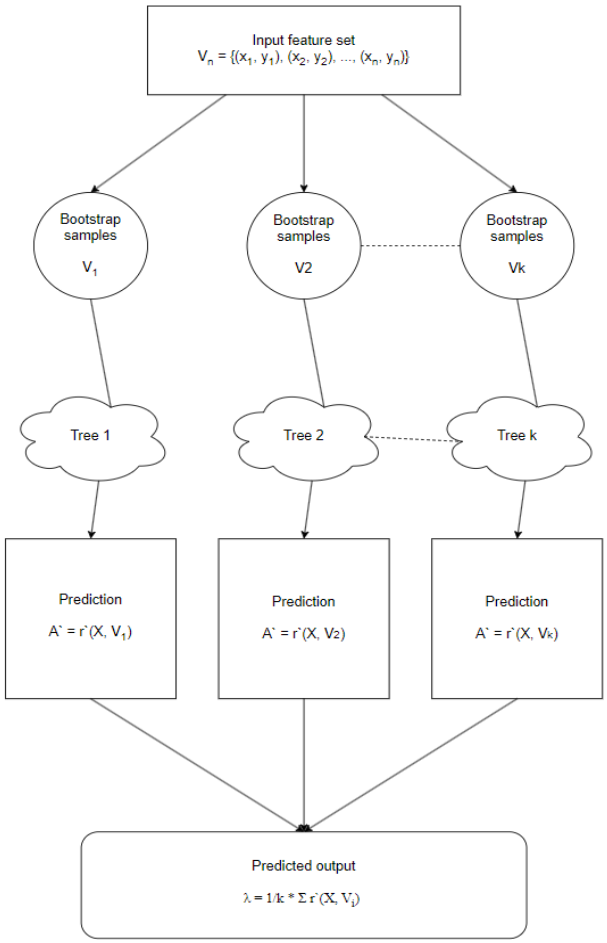



Figure 1 – Block diagram of the Random Forest algorithm

The next step, after identifying the most vulnerable part of the population, is to model the spread of TB in the regions of Ukraine. To perform this part of the task, it was decided to use the SIR model. The SIR model is a simple cellular automata model used to simulate the spread of infection. It divides people into three groups:

– Susceptible (*S*): People who can contract an infection.

– Infected (*I*): People who are infected with the infection and can transmit it to others.

– Recovered (*R*): People who have recovered from an infection and can no longer be infected.

The SIR model works as follows:

– Infected people transmit the infection to susceptible people with a certain probability.

– Susceptible people who are exposed to the infection become infected.

– Infected people eventually recover.

Fig. 2 shows an image of a compartmental diagram of the SIR model, where $N$, $S_0$, $I_0$, $R_0$ are the total population, the initial value of susceptible people , the initial value of infected people, and the initial value of people who can no longer get sick, respectively, β is a symbol is the symbol for the infection rate (describes the probability that a

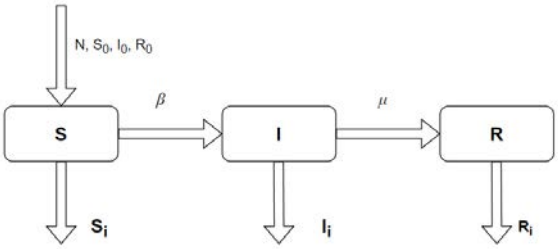



Figure 2 – Compartmental diagram of the SIR model

susceptible individual will become infected through contact with an infected individual), μ is the symbol for the mortality/cure rate (describes the probability that infected individuals will die of the disease or recover without the possibility of contracting the disease again).

The values of $S_i$, $I_i$, and $R_i$ are calculated from the differential equations at each iteration of the simulation as follows, respectively [12, 15]:

$$S_i = \frac{-\beta}{N * S_{i-1} * I_{i-1}},$$

$$I_i = \frac{\beta}{N * S_{i-1} * I_{i-1} * (I - \mu)},$$

$$R_i = I * \mu.$$

Here is a pseudo-code of the algorithm:

```
Algorithm SIR: pseudocode
1:   function MAIN
2:   set t, fixed = (p^(fix), t), bounds = range for
     p^(cal)
3:   for run = 0 to Number_of runs parallel do
4:     call single run (bounds, fixed)
5:   end for
6:   read and assembles results from single runs
7:   end function
8:   function SINGLE_RUN(bounds, fixed)
9:   call differential_evolution (Objective func-
     tion, bounds, fixed)
10:  save results of a single run
11:  end function
12:  function OBJECTIVE_FUNCTION(p, t)
13:  s = SIRmodel (p)
14:  objfun=O_y(s),t(p)
15:  return objfun
16:  end function
17:  function SIRMODEL(p)
18:  initialize S_nini
19:  for n = n_ini + 1 to n_ini + N^(mod) - 1 do
20:    compute s_n from s_n-1 by (10)
21:  end for
22:  return s
23:  end function
```

Before dividing the data into training and testing subsets, it is necessary to filter the data. To do this, it is necessary to delete the tables that will not be used in the analysis, for example “Block 4. Information about TB facilities”, “Block 5. TB treatment”, “Block 6. Bed capacity of TB facilities”, and other separate tables. Other tables should be combined by years and regions, and the data should be cleaned (filling in gaps if any, normalisation, standardisation, etc.). Examples of table blocks divided by the information they store are shown in Fig. 3.


DOI 10.15588/1607-3274-2024-4-4

| | |
|---|---|
| Block 3. Tuberculosis / HIV infection | |
| Table 44 | Incidence of tuberculosis in combination with AIDS (new cases + relapses) |
| Table 45 | Registration of HIV-positive persons with tuberculosis |
| Table 46 | Tuberculosis patients died from the disease caused by AIDS |
| Table 47 | Prevalence of all forms of active tuberculosis in combination with the disease caused by HIV |
| Block 4. Availability of doctors in anti-tuberculosis institutions | |
| Table 48 | Availability of phthisiologist doctors in the institutions of the Ministry of Health of Ukraine |
| Table 49 | Medical positions in medical and preventive institutions of the Ministry of Health of Ukraine, 2021 |
| Block 6. Bed fund of anti-tuberculosis institutions | |
| Table 69 | The network of anti-tuberculosis healthcare institutions of the Ministry of Health of Ukraine |
| Table 70 | Availability of hospital beds for tuberculosis patients in health care institutions of the Ministry of Health of Ukraine |
| Table 71 | Indicators of the use of the bed fund of anti-tuberculosis health care institutions of the Ministry of Health of Ukraine, 2021 |
| Table 72 | Hospital and sanatorium care for tuberculosis patients by the territorial location of health care institutions of the Ministry of Health of Ukraine |

Figure 3 – Examples of table blocks from the data set

– To perform the analysis, it is necessary to design 2 data sets. The first set, which will be used to predict the age and sex categories of people who are most vulnerable to further infection using the Random Forest algorithm, will consist of the following columns: AgeCategory, Year, Sex, TotalPopulation, Infected, Dead. An example from this dataset is shown in Fig. 4.

| AgeCategory | Sex | Year | TotalPopulation | Infected | Dead |
|---|---|---|---|---|---|
| 0–1 | M | 2007 | 46650000 | 15 | 2 |
| 1–4 | M | 2007 | 46650000 | 113 | 0 |
| 5–9 | M | 2007 | 46650000 | 107 | 1 |
| 10–14 | M | 2007 | 46650000 | 99 | 1 |
| 15–17 | M | 2007 | 46650000 | 300 | 1 |
| 18–24 | M | 2007 | 46650000 | 2445 | 162 |
| 25–34 | M | 2007 | 46650000 | 6195 | 1365 |
| 35–44 | M | 2007 | 46650000 | 6310 | 2405 |
| 45–54 | M | 2007 | 46650000 | 5958 | 2890 |
| 55–64 | M | 2007 | 46650000 | 2831 | 1195 |
| 65–100 | M | 2007 | 46650000 | 1895 | 586 |
| 0–1 | F | 2007 | 46650000 | 7 | 1 |
| 1–4 | F | 2007 | 46650000 | 83 | 3 |
| 5–9 | F | 2007 | 46650000 | 90 | 0 |
| 10–14 | F | 2007 | 46650000 | 107 | 1 |
| 15–17 | F | 2007 | 46650000 | 303 | 1 |
| 18–24 | F | 2007 | 46650000 | 1707 | 91 |
| 25–34 | F | 2007 | 46650000 | 2821 | 464 |
| 35–44 | F | 2007 | 46650000 | 2118 | 512 |
| 45–54 | F | 2007 | 46650000 | 1482 | 449 |
| 55–64 | F | 2007 | 46650000 | 806 | 171 |
| 65–100 | F | 2007 | 46650000 | 1303 | 205 |
| 65–100 | | | | | |

Figure 4 – An example of a designed data set for the Random Forest algorithm

Table 2 – Description of data set elements for the Random Forest algorithm

| Feature name | Description |
|---|---|
| AgeCategory | Age categories represented in the original dataset from 0 to 100 years |
| Sex | Sex |
| Year | The year, used for reference |
| TotalPopulation | Country population |
| Infected | Amount of infected people |
| Dead | Amount of dead people |

Table 2 presents a description of the features that make up the data set for the Random Forest algorithm.

For the dataset that will be used by the SIR algorithm to model the spread of the disease, we will design a dataset consisting of the following columns: Year, InfectedTB, DeadTB, Region, TotalPopulation, RecoveredTB. An example of the generated data is shown in Fig. 5.

| Year | InfectedTB | DeadTB | Region | TotalPopulation | RecoveredTB |
|---|---|---|---|---|---|
| 2022 | 405 | 93 | Volyn | 1018628 | 150 |
| 2022 | 2450 | 340 | Dnipropetrovsk | 3093176 | 1013 |
| 2022 | 133 | 13 | Donetsk | 1883713 | 58 |
| 2022 | 491 | 98 | Zhytomyr | 1179801 | 189 |
| 2022 | 880 | 136 | Zakarpattia | 1241643 | 358 |
| 2022 | 595 | 161 | Zaporizhzhia | 1637673 | 209 |
| 2022 | 378 | 44 | Ivano-Frankivsk | 1349096 | 161 |
| 2022 | 857 | 129 | Kyiv | 1789300 | 350 |
| 2022 | 607 | 100 | Kirovohrad | 897297 | 244 |
| 2022 | 13 | 78 | Luhansk | 666801 | –32 |
| 2022 | 846 | 195 | Lviv | 2459763 | 313 |
| 2022 | 952 | 85 | Mykolaiv | 1091106 | 417 |
| 2022 | 2927 | 249 | Odessa | 2340332 | 1286 |
| 2022 | 700 | 70 | Poltava | 1344445 | 303 |
| 2022 | 530 | 67 | Rivne | 1140724 | 223 |
| 2022 | 271 | 88 | Sumy | 1033580 | 88 |
| 2022 | 280 | 29 | Ternopil | 1018462 | 121 |
| 2022 | 965 | 181 | Kharkiv | 2583325 | 377 |
| 2022 | 476 | 100 | Kherson | 1000166 | 181 |
| 2022 | 441 | 50 | Khmelnytskyi | 1225666 | 188 |
| 2022 | 419 | 95 | Cherkasy | 1157115 | 156 |
| 2022 | 406 | 42 | Chernivtsi | 887392 | 175 |
| 2022 | 469 | 78 | Chernihiv | 950773 | 188 |

Figure 5 – Example of a designed dataset for the SIR model

Table 3 – Description of data set features for the SIR model algorithm

| Feature name | Description |
|---|---|
| Year | Year, used for reference |
| Region | Region of Ukraine (not included city of Kyiv) |
| RecoveredTB | Amount of recovered people |

Table 3 provides a description of the features included in the dataset for the SIR algorithm.


DOI 10.15588/1607-3274-2024-4-4

## 4 EXPERIMENTS

Conducting experiments for the topic “Modelling the spread of tuberculosis in the regions of Ukraine for 2024 with the identification of the most vulnerable groups by gender and age” is of great importance, as it allows us to understand how accurately the selected models or algorithms have worked, namely, to understand how accurately the Random Forest model predicts the number of infections in the upcoming years, which we can use to make conclusions about the most vulnerable part of the population. In addition, the experiments make it possible to visually see the modelled dynamics of the spread of the disease in each of the country’s regions using the SIR cellular automata model.

The software implementation was done using Python, a high-level general purpose programming language with a simple and readable syntax. It has a large number of libraries for a variety of tasks, making it a very powerful tool for solving problems in research, data analysis, and machine learning. Python’s advantages include ease of learning, broad community support, and ease of use. In particular, the following libraries and tools were used to develop the application:

– Pandas is a data processing and analysis library that provides data structures and functions for working with them. It allows you to easily perform operations on large data sets, such as reading, writing, filtering, and aggregating data.

– NumPy is a library for scientific computing in Python. It provides support for arrays and mathematical functions, allowing you to easily perform calculations on numerical data.

– Scikit-learn is a machine learning library for Python. It contains implementations of many machine learning algorithms, such as classification, regression, clustering, and others, as well as tools for estimating and fitting model parameters.

– Seaborn is a Python data visualisation library based on the Matplotlib library. It provides high-level functions for creating attractive and informative graphs and charts.

– Matplotlib is a graphing and data visualisation library in Python. It allows you to create various types of graphs, such as line, pie, and bar charts, and customise their appearance.

– Scipy is a Python library for scientific and technical computing. It contains implementations of many algorithms for numerical computation, optimisation, signal processing, and other functions for working with scientific data.

The dataset for the Random Forest algorithm described in the previous section, an example of which can be seen in Fig. 6, is stored in .csv format, so you need to write a function that reads data from the file and returns it in the pandas Data frame format:

```
Function readDataFrame(file path):
1: initialize an empty DataFrame called df_rf
2: read the contents of the file using
   "pd.read_csv()"
3: add the contents of the read file to df_rf
4: return df_rf
```

| № | Age-Category | Sex | Year | TotalPopu-lation | In-fected | Dead |
|---|---|---|---|---|---|---|
| 0 | 0–1 | M | 2007 | 46650000 | 15 | 2 |
| 1 | 1–4 | M | 2007 | 46650000 | 113 | 0 |
| 2 | 5–9 | M | 2007 | 46650000 | 107 | 1 |
| 3 | 10–14 | M | 2007 | 46650000 | 99 | 1 |
| 4 | 15–17 | M | 2007 | 46650000 | 300 | 1 |
| … | … | … | … | … | … | … |
| 347 | 25–34 | F | 2022 | 41167000 | 753 | 63 |
| 348 | 35–44 | F | 2022 | 41167000 | 1119 | 134 |
| 349 | 45–54 | F | 2022 | 41167000 | 952 | 121 |
| 350 | 55–64 | F | 2022 | 41167000 | 644 | 79 |
| 351 | 65–100 | F | 2022 | 41167000 | 787 | 77 |

Figure 6 – Data set in the pandas DataFrame format

Fig. 7 shows an example of data from a loaded dataset. After reading, you need to check the data types that are stored in the set. The result of the check can be seen in Fig. 8.

For the model to work, you must first clean the data, check for spaces, and convert it to the required types, after which the data set looks like this:

| | |
|---|---|
| AgeCategory | object |
| Sex | object |
| Year | int64 |
| TotalPopulation | object |
| Infected | object |
| Dead | int64 |
| dtype | object |

Figure 7 – Data types of the downloaded dataset

| № | AgeCa tegory | Sex | Year | TotalPopul ation | Infected | Dead |
|---|---|---|---|---|---|---|
| 0 | 0–1 | M | 2007 | 46650000 | 15 | 2 |
| 1 | 1–4 | M | 2007 | 46650000 | 113 | 0 |
| 2 | 5–9 | M | 2007 | 46650000 | 107 | 1 |
| 3 | 10–14 | M | 2007 | 46650000 | 99 | 1 |
| 4 | 15–17 | M | 2007 | 46650000 | 300 | 1 |
| … | … | … | … | … | … | … |
| 347 | 25–34 | F | 2022 | 41167000 | 753 | 63 |
| 348 | 35–44 | F | 2022 | 41167000 | 1119 | 134 |
| 349 | 45–54 | F | 2022 | 41167000 | 952 | 121 |
| 350 | 55–64 | F | 2022 | 41167000 | 644 | 79 |
| 351 | 65–100 | F | 2022 | 41167000 | 787 | 77 |

Figure 8 – Data after cleaning

Fig. 9 shows an example of data after data cleaning operations.

| | |
|---|---|
| AgeCategory | object |
| Sex | object |
| Year | int64 |
| TotalPopulation | int64 |
| Infected | int64 |
| Dead | int64 |
| dtype | object |

Figure 9 – Data types after cleaning the dataset


DOI 10.15588/1607-3274-2024-4-4

Fig. 10 shows the types of data from the dataset for the Random Forest algorithm after data cleaning operations.

| | |
|---|---|
| AgeCategory | 0 |
| Sex | 0 |
| Year | 0 |
| TotalPopulation | 0 |
| Infected | 0 |
| Dead | 0 |
| dtype | int64 |

Figure 10 – Checking for gaps in the data set

As can be seen from Fig. 10, the data in the dataset has no gaps, so we can proceed to normalise the required data columns, namely TotalPopulation, Infected and Dead, and perform label encoding for the Sex and AgeCategory columns. Normalisation will be performed using the following function:

```
   Function min_max_normalize(dataset column):
1: We look for the minimum value in the column
   and set it in the col_min variable
2: We look for the maximum value in the column
   and set it in the col_max variable
3: We calculate the updated value of the column
   elements according to the formula (column
   value – col_min) / (col_max – col_min)
4: We return the column with the updated values
```

The label encoding operation will be performed using the LabelEncoder function, which we import from the sklearn.preprocessing library.

After performing the operations, mentioned above, the new view of the dataset for the Random Forest algorithm can be seen in Fig. 11.

| № | Age-Category | Sex | Year | Total-Population | Infected | Dead |
|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 2007 | 1.0 | 0.00146 | 0.00068 |
| 1 | 1 | 1 | 2007 | 1.0 | 0.01574 | 0.00000 |
| 2 | 8 | 1 | 2007 | 1.0 | 0.01487 | 0.00034 |
| 3 | 2 | 1 | 2007 | 1.0 | 0.0137 | 0.00034 |
| 4 | 3 | 1 | 2007 | 1.0 | 0.043 | 0.00034 |
| … | … | … | … | … | … | … |
| 347 | 5 | 0 | 2022 | 0.0 | 0.109 | 0.02152 |
| 348 | 6 | 0 | 2022 | 0.0 | 0.162 | 0.04578 |
| 349 | 7 | 0 | 2022 | 0.0 | 0.13805 | 0.04134 |
| 350 | 9 | 0 | 2022 | 0.0 | 0.0932 | 0.027 |
| 351 | 10 | 0 | 2022 | 0.0 | 0.114 | 0.02631 |

Figure 11 – View of the dataset after its preparation

To better understand the relationship between the data in the dataset, let's illustrate the correlation table shown in Fig. 12.

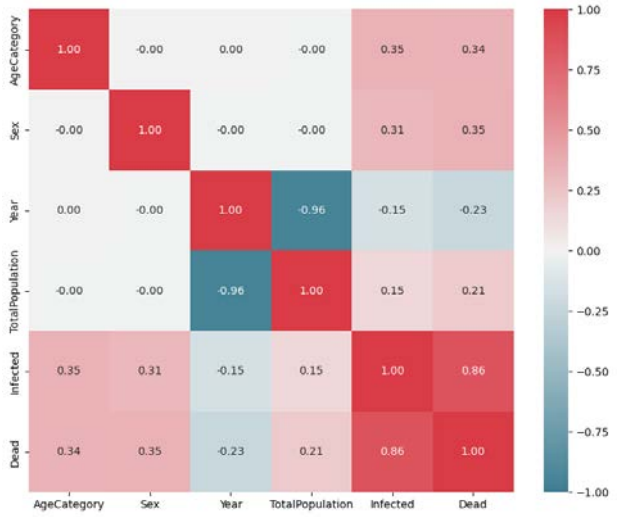



Figure 12 – Correlation table of the dataset

Cells in the correlation table with values from 0.7 to 1.0 will be considered strong correlations, while the medium correlation will be considered to be cells with values between 0.3 and 0.69. As we can see from Fig. 12, the features Infected and Dead, AgeCategory and Infected, Sex and Infected have an average correlation, while the features TotalPopulation and Year have a strong inverse correlation.

Before starting the algorithm, it is necessary to select the target variable, which will be the Infected feature, and all other features will be used for prediction. Also, it is necessary to divide the data set into training and testing, which will be done in the ratio of 7:3.

Finally, we can move on to the model that will be used for prediction. Given that the target feature is a continuous value, we used the RandomForestRegressor algorithm taken from the Python library sklearn.ensemble, where the number of trees to be used is set to 100 as a parameter.

Once the model is trained, we run the test, and evaluate the model's performance, using the metrics of root mean square error and $R^2$ estimation. The following results are obtained:

```
Mean Squared Error (MSE): 0.0028
R-squared: 0.92
```

As you can see, the value of the mean squared error is quite small, while the value of the $R^2$ estimate is quite high, with the highest possible value of this estimate being 1.

We also test the model by checking which population group, by age and gender, is most vulnerable to infection. We have the following results:

```
Based on the model, the group with the
highest predicted number of deaths in the
year 2024 is:
Age Category: 35–44
Sex: M
```


DOI 10.15588/1607-3274-2024-4-4

Based on these results, we can conclude that men aged 35 to 44 will have the highest number of infections in the next year. Also, to better understand the situation of infection for other categories of sex and age, we visualise the graph of the predicted spread of the disease among them, which is shown in Fig. 13.

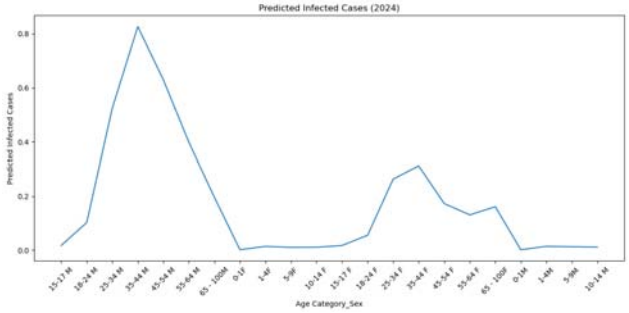



Figure 13 – Statistics of subsequent diseases for all categories of gender and age

From Fig. 13, we can see that the age group of men aged 35–44 is the most susceptible to infection, while the second place is occupied by the group of women of the same age category. The values for other age groups for both sexes were also illustrated in Fig. 13.

To model the spread of the disease in each of the 24 regions of Ukraine and in the city of Kyiv, it is necessary to use the data set, an example of which can be seen in Fig. 14. From the example of this data set, we have information on the number of infected, the number of recovered, the number of deaths, the year, the name of the region, and the population of the respective region.

Let's load the data set into the program in the pandas DataFrame format using the *readDataFrame* function.

| № | Year | Infect-edTB | Dead TB | Region | Total-Popula-tion | Recov-eredTB |
|---|---|---|---|---|---|---|
| 411 | 2022 | 441 | 50 | Khmel-nytskyi | 1225666 | 188 |
| 412 | 2022 | 419 | 95 | Cherkasy | 1157115 | 156 |
| 413 | 2022 | 406 | 42 | Chernivtsi | 887392 | 175 |
| 414 | 2022 | 469 | 78 | Chernihiv | 950773 | 188 |
| 415 | 2022 | 644 | 102 | KyivCity | 2910994 | 261 |

Figure 14 – Example of a data set in the DataFrame format

Fig. 14 shows an example of data from a dataset for the SIR algorithm after conversion to the DataFrame format. After reading, it is necessary to check the data types stored in the set, which results in the following result:

| | |
|---|---|
| Year | int64 |
| InfectedTB | object |
| DeadTB | int64 |
| Region | object |
| TotalPopulation | object |
| RecoveredTB | int64 |
| dtype | object |

Figure 15 – Data types of the loaded dataset before cleaning

From Fig. 15, we can see that some numeric features that should represent numeric values do not have the required data type set, for which it is necessary to perform preliminary data cleaning and convert the required features to the required data types, and then we have:

| | |
|---|---|
| Year | int64 |
| InfectedTB | int64 |
| DeadTB | int64 |
| Region | string |
| TotalPopulation | int64 |
| RecoveredTB | int64 |
| dtype | object |

Figure 16 – Data types of the loaded dataset after cleaning

Fig. 16 shows a description of the data types for the features of the dataset for the SIR algorithm after the preparatory operations. Also, let's check the data for completeness, that is, whether it has any empty fields. The result of this check is shown in Fig. 17.

| | |
|---|---|
| Year | 0 |
| InfectedTB | 0 |
| DeadTB | 0 |
| Region | 0 |
| TotalPopulation | 0 |
| RecoveredTB | 0 |
| dtype | int64 |

Figure 17 – Checking the dataset for completeness

As we can see from Fig. 17, the data has no voids. Also, to better understand the relationships between the features in the dataset, let's illustrate the correlation table shown in Fig. 18.


Figure 18 – Correlation table for a data set

Describing the correlation table for the Random Forest dataset shown in Fig. 18, we defined the terms strong and medium correlation, so that the features InfectedTB and RecoveredTB, DeadTB and RecoveredTB, InfectedTB and DeadTB have a strong correlation, while almost all features at the intersection with the feature Year have a strong inverse correlation.


DOI 10.15588/1607-3274-2024-4-4

Before starting the SIR algorithm, it is necessary to describe the initial values of the parameters *S*, *I*, *R*, which stand for the number of susceptible people, the number of infected people, and the number of people who can no longer get sick. In the latter category, we included both people who have recovered and people who have already died from the disease. Among the parameters, there is also one that determines the number of days during which the disease spread will be modelled. In our case, it is set to 700 days. Also, for the algorithm to work, it is necessary to set the parameters β and μ, which act as coefficients in calculating the number of people moving from state *S* to state *I* (transmission coefficient or the rate at which susceptible people become infected when they come into contact with infected people. It indicates the rate at which the disease spreads in the population), and from state *I* to state *R* (the rate of recovery or the rate at which infected individuals recover from the disease and become immune), respectively. Thus, for the above parameters, which are necessary for the algorithm to work, we have set the following values:

*S* = TotalPopulation – InfectedTB – RecoveredTB – DeadTB;

*I* = InfectedTB;

*R* = number of recovered TB cases + number of dead TB cases;

β = 4/10 (which means that every 10 days 4 people become infected);

**μ** = 1/10 (which means that every 10 days one person recovers);

Finally, when the algorithm is finished, we have the following results of modelling the number of eligible, infected, recovered or dead people for some regions in Fig. 19–21:

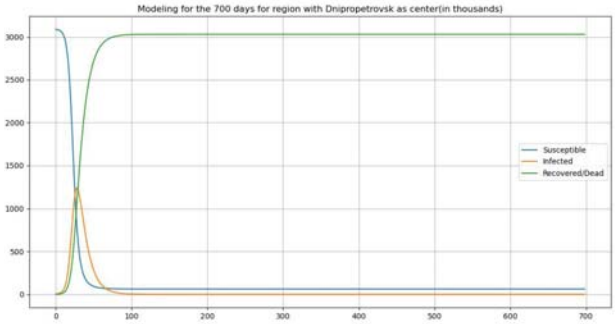



Figure 19 – Modelling results for Dnipro region

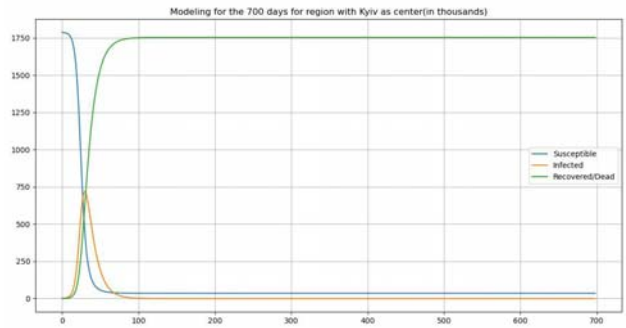



Figure 20 – Modelling results for Kyiv region

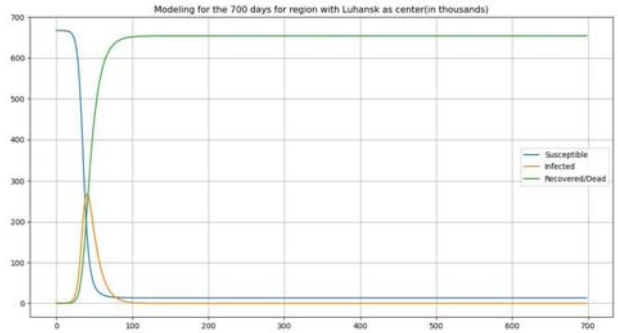



Figure 21 – Modelling results for Luhansk region

To consider examples of modelling results shown in Fig. 19–21, we have selected 3 regions of Ukraine that differ in population size, namely Dnipro region(population of almost 3 million people), Kyiv region(population of almost 1.8 million people) and Luhansk region(population of almost 0.7 million people). As we can see, regardless of the initial population (or people eligible for the disease), the largest increase in infected persons occurs in the first 100 days of the simulation, after which the number of people who can be infected and the number of infected persons drop to 0, moving to the recovered/dead state, the number of which is close to the initial number of people who can be infected.

To better understand the state of the disease in each of the oblasts in comparison to other oblasts, let us illustrate the maximum percentage of infected people in each of the oblasts for the entire modelling period relative to the total population in the oblast in Fig. 22.

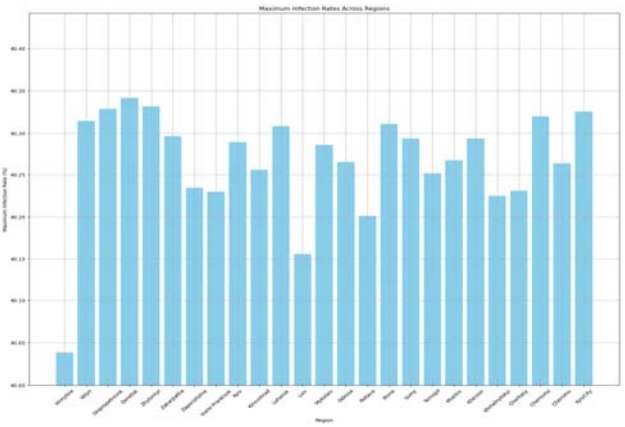


Figure 22 – Comparative histogram of the maximum number of infected people in relation to the total population of the region

From Fig. 22 shows that the maximum number of infected individuals relative to the total population of the region was approximately the same for the entire modelling period, fluctuating around 40%. The region with the lowest value of the maximum number of infected individuals is Vinnytsia, and the highest is Donetsk.


DOI 10.15588/1607-3274-2024-4-4

## 5 RESULTS

Chapter 4 described the algorithms and demonstrated their results. It is important not only to evaluate the results visually, but also with the help of metrics. As we can see from Section 4.2, the MSE and $R^2$ metrics were evaluated for the Random Forest algorithm. Under ideal conditions, the value of the MSE metric should be close to 0, while the value of the $R^2$ metric should be close to one.

Table 3 – Metrics values for the results of the Random Forest algorithm

| Metric | Volue |
|---|---|
| MSE | 0.0028 |
| $R^2$ | 0.9204 |

As we can see from Table 3, the actual values of the metrics are quite close to the values of the metrics under ideal conditions, and the prediction provided by this algorithm is quite accurate, as can be seen in the statistical analysis of the data set.

The operation of the SIR algorithm was also described in section 4.3, after which, by calculating the results separately for each of the regions and visualising the maximum number of infected individuals relative to the total population of the region for the entire modelling period in Fig. 4.17, we were able to verify the lack of accuracy of the modelling of the disease spread, since the average value for all regions in Fig. 25 reaches 39%, which cannot be true in reality. Such inaccuracies in the results can be explained by the fact that risk factors for the disease, the possibility of moving from a state of recovery to a state of infection, and some others are not taken into account. Also, the parameters that are taken into account when transitioning between the states of healthy-infected and infected-recovered/dead, are an important part in this algorithm. These parameters need to be better tuned, which requires a larger scientific base.

In general, comparing the selected algorithms with those chosen by other researchers in related topics, we can clearly see that the selected models are architecturally the simplest in their field, and of course, their simplicity takes away from their high accuracy of results.

For further analysis of this topic, it is worth choosing more structurally complex algorithms, the accuracy of which will be higher, namely recurrent neural networks, or forecasting using time series.

## 6 DISCUSSION

Conducted research aimed at studying the spread, modeling of tuberculosis and forecasting the most vulnerable social groups of the population. During the analysis of the literature, the relevance of the problem was revealed and methods for modeling the disease were chosen.

A comparative analysis of other methods used in previous studies was conducted to determine the most effective methods in this context. The use of the Random Forest method has demonstrated sufficient effectiveness in predicting vulnerable social groups of the population in the context of the spread of tuberculosis.

The SIR algorithm proved to be less effective in modeling the spread of the disease due to its shortcomings identified during the study. It is necessary to consider more complex algorithms to obtain more accurate results in predicting the spread of tuberculosis.

For the further development of the research, it is recommended to take into account additional factors affecting the spread of the disease.

It is recommended to carry out a simulation of the spread of tuberculosis among the population of Ukraine for a better understanding of further actions in the fight against this disease.

## CONCLUSIONS

Comparing the selected algorithms with other studies in related topics, one can understand that they are architecturally the simplest in their field.

The simplicity of these models takes away the ability to provide high accuracy results, but for further analysis of the topic, it is recommended to consider more architecturally complex algorithms, such as recursive neural networks or time series forecasting, which can provide even greater accuracy of predictions.

The study compared in detail the effectiveness of two different modeling approaches – machine learning (Random Forest) and epidemiological model (SIR). The Random Forest method was found to provide higher prediction accuracy, which is important for further research and practical work in the field of healthcare. A preliminary data analysis was carried out with the selection of the most suitable sets for the study, which increases the relevance and accuracy of the results obtained.

**Practical significance of the results**: Conclusions about the effectiveness of the Random Forest method can be used to create programs for predicting and controlling the spread of tuberculosis, which is an urgent task for the health care system.

**Recommendations for further research:** Based on the obtained results, recommendations have been developed regarding the use of more complex algorithms and consideration of additional factors that may affect the accuracy of models, which contributes to the development of a scientific approach to the study of the spread of diseases.

These aspects ensure the **scientific novelty** of the work and emphasize its significance in the context of tuberculosis research and forecasting of vulnerable social groups.

**Prospects for further research** are to study the proposed algorithms for a wide class of practical problems.

## ACKNOWLEDGEMENTS

The study was created within research topic “Methods and means of artificial intelligence to prevent the spread of tuberculosis in wartime” (№0124U000660), which is carried out at the Department of Artificial Intelligence


DOI 10.15588/1607-3274-2024-4-4

Systems of the Institute of Computer Sciences and Information of technologies of the National University “Lviv Polytechnic”.

## REFERENCES

1. Duko B., Bedaso A., Ayano G. The prevalence of depression among patients with tuberculosis: a systematic review and meta-analysis, *Ann Gen Psychiatry*, 2020, Vol. 19(1), P. 1. Mode of access: https://doi.org/10.1371/journal.pone.0227472.
2. [Ruiz-Grosso P., Cachay R., De La Flor A. et al.] Association between tuberculosis and depression on negative outcomes of tuberculosis treatment: a systematic review and meta-analysis, *PLoS ONE*, 2020, Vol. 15(1), P. 1. Mode of access: https://doi.org/10.1371/journal.pone.0227472.
3. Cohen A., Mathiasen V. D., Schön T. et al. The global prevalence of latent tuberculosis: a systematic review and meta-analysis, *Eur Respir J*, 2019, Vol. 54(3), P. 1. Mode of access: https://doi.org/10.1183/13993003.00655-2019.
4. Kiazyk S., Ball T. Tuberculosis (TB): latent tuberculosis infection: an overview, *Canada Commun Dis Rep.*, 2017, Vol. 43(3–4), P. 62. Mode of access: https://doi.org/10.14745/ccdr.v43i34a01.
5. World Health Organization. Annual Report of Tuberculosis. Annual Global TB Report of WHO [Electronic resource]. Access mode: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 (access date: 11.05.2024). Title from the screen.
6. Abascal E., Pérez-Lago L., Martínez-Lirola M. et al. Whole genome sequencing-based analysis of tuberculosis (TB) in migrants: Rapid tools for crossborder surveillance and to distinguish between recent transmission in the host country and new importations, *Eurosurveillance*, 2018, Vol. 24(4), P. 1800005. Mode of access: https://doi.org/10.2807/1560-7917.ES.2019.24.4.1800005.
7. De Beer J. L., Ködmön C., Van der Werf M. J. et al. Molecular surveillance of multi- and extensively drug-resistant tuberculosis transmission in the European Union from 2003 to 2011, *Eurosurveillance*, 2014, Vol. 19(11), P. 20742. Mode of access: https://doi.org/10.2807/1560-7917.ES2014.19.11.20742.
8. Dohál M., Dvořáková V., Šperková M. et al. Whole genome sequencing of multidrug-resistant Mycobacterium tuberculosis isolates collected in the Czech Republic, 2005–2020, *Sci Rep.*, 2022, Vol. 12(1), pp. 1–10. Mode of access: https://doi.org/10.1038/s41598-022-11287-5.
9. Dohál M., Dvořáková V., Šperková M. et al. Antituberculosis drug resistance in Slovakia, 2018–2019: the first whole-genome epidemiological study, *J Clin Tuberc Other Mycobact Dis*, 2022, Vol. 26, P. 100292. Mode of access: https://doi.org/10.1016/j.jctube.2021.100292.
10. Pavlenko E., Barbova A., Hovhannesyan A. et al. Alarming levels of multidrug-resistant tuberculosis in Ukraine: results from the first national survey, *Int J Tuberc Lung Dis*, 2018, Vol. 22(2), pp. 197–205. Mode of access: https://doi.org/10.5588/ijtld.17.0254.
11. Vyklyuk Ya. Nevinskyi D., Boyko N. GeoCity – a New Dynamic-Spatial Model of Urban Ecosystem, *J. Geogr. Inst. Cvijic*, 2023, Vol. 73(2), pp. 187–203. Mode of access: https://doi.org/10.2298/IJGI2302187V.
12. Zignol M., Cabibbe A. M., Dean A. S. et al. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study, *Lancet Infect Dis*, 2018, Vol. 18(6), pp. 675–683. Mode of access: https://doi.org/10.1016/S1473-3099(18)30073-2.
13. Paul L. K., Nchasi G., Bulimbe D. B. et al. Public health concerns about Tuberculosis caused by Russia/Ukraine conflict, *Health Sci Rep*, 2023, Vol.6(4), P. 1. Mode of access: https://doi.org/10.1002/hsr2.1218.
14. Aldridge R. W., Zenner D., White P. J. et al. Tuberculosis in migrants moving from high-incidence to low-incidence countries: a population-based cohort study of 519,955 migrants screened before entry to England, Wales, and Northern Ireland, *Lancet*, 2016, Vol. 388(10059), P. 2510. Mode of access: https://doi.org/10.1016/S0140-6736(16)31008-X.
15. Merker M., Blin C., Mona S. et al. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage, *Nat Genet*, 2015, Vol. 47(3), pp. 242–249. Mode of access: https://doi.org/10.1038/ng.3195.
16. Daum L. T., Konstantynovska O. S., Solodiankin O. S. et al. Next-generation sequencing for characterizing drug resistance-conferring mycobacterium tuberculosis genes from clinical isolates in the Ukraine, *J Clin Microbiol*, 2018, Vol. 56(6), P. 1. Mode of access: https://doi.org/10.1128/JCM.00009-18.
17. Huang C. C., Chu A. L., Becerra M. C. et al. Mycobacterium tuberculosis Beijing lineage and risk for tuberculosis in child household contacts, *Peru. Emerg Infect Dis*, 2020, Vol. 26(3), P. 568. Mode of access: https://doi.org/10.3201/eid2603.191314.
18. Vyklyuk Y., Semianiv I., Nevinskyi D. et al. Applying geospatial multi-agent system to model various aspects of tuberculosis transmission, *New Microbes and New Infections*, 2024, Vol. 59, P. 101417. Mode of access: https://doi.org/10.1016/j.nmni.2024.101417.
19. Stucki D., Ballif M., Egger M. et al. Standard genotyping overestimates transmission of mycobacterium tuberculosis among immigrants in a low-incidence country, *J Clin Microbiol*, 2016, Vol. 54(7), pp. 1862–70. Mode of access: https://doi.org/10.1128/JCM.00126-16.
20. Jackson S., Kabir Z., Comiskey C. Effects of migration on tuberculosis epidemiological indicators in low and medium tuberculosis incidence countries: a systematic review, *J Clin Tuberc Other Mycobact Dis*, 2021, Vol. 23, pp. 2405–5794. Mode of access: https://doi.org/10.1016/j.jctube.2021.100225.

Received 20.08.2024.
Accepted 24.10.2024.


DOI 10.15588/1607-3274-2024-4-4

УДК 004.94

# МОДЕЛЮВАННЯ ПОШИРЕННЯ ТУБЕРКУЛЬОЗУ ЗА РЕГІОНАМИ В УКРАЇНІ

**Бойко Н. І.** – канд. економ. наук, доцент, доцент кафедри Систем штучного інтелекту, Національний університет «Львівська політехніка», Львів, Україна.

**Работягов Д. С.** – студент кафедри Системи штучного інтелекту, Національний університет «Львівська політехніка», Львів, Україна.

## АНОТАЦІЯ

**Актуальність.** Моделювання поширення туберкульозу на території України є особливо актуальним у зв'язку зі зростанням числа випадків захворювання, зокрема у 2023 році.

**Мета роботи** є вирішення задач моделювання шляхом застосування сучасних методів машинного навчання та аналізу даних для побудови прогностичних моделей поширення туберкульозу на регіональному рівні.

**Метод.** Для моделювання поширення туберкульозу на регіональному рівні в Україні пропонується використовувати кілька підходів, таких як SIR модель, клітинні автомати та Random Forest. Кожен з цих методів має свої унікальні переваги та може забезпечити детальніше розуміння динаміки поширення захворювання. SIR модель (Susceptible-Infectious-Recovered) є класичною епідеміологічною моделлю, яка описує розповсюдження інфекційних захворювань у популяції. Модель передбачає три групи населення: *S* (Susceptible) – сприйнятливі до інфекції; *I* (Infectious) – інфіковані та здатні передавати інфекцію; *R* (Recovered) – ті, хто одужав та отримав імунітет. Клітинні автомати є дискретною моделлю, що використовує решітку клітин для моделювання просторово-часових процесів. Кожна клітина може перебувати у різних станах (наприклад, здорова, інфікована, імунна) та змінювати свій стан залежно від стану сусідніх клітин. Random Forest є методом машинного навчання, що використовує ансамбль дерев рішень для класифікації або регресії. Цей метод може бути застосований для прогнозування поширення туберкульозу на основі великої кількості вхідних параметрів. Використання цих методів дозволить провести глибокий аналіз та отримати комплексні результати щодо поширення туберкульозу на регіональному рівні в Україні. Це, в свою чергу, сприятиме розробці ефективних стратегій боротьби з хворобою та покращенню здоров'я населення.

**Результати.** Були детально описані та проаналізовані результати застосування методів Random Forest і SIR. Для Random Forest були оцінені метрики MSE та $R^2$, що показали високу точність передбачень. У випадку моделювання алгоритмом SIR, за допомогою візуальної оцінки результатів, було виявлено недостатню точність, що обумовлено недоліками моделі. Порівнюючи обрані методи з іншими дослідженнями, було зроблено висновок про необхідність розгляду більш складних алгоритмів для отримання більш точних результатів.

**Висновки.** На основі результатів дослідження можна зробити висновок про достатню ефективність методу Random Forest для та прогнозування уразливих соціальних груп населення та слабку ефективність алгоритму SIR для моделювання поширення туберкульозу. Для подальшого розвитку дослідження рекомендується розгляд більш складних алгоритмів та врахування додаткових факторів, що впливають на поширення захворювання. Крім того, для кращого розуміння подальших дій для поротьби з хворобою, доцільно буде провести симуляцію поширення туберкульозу серед населення України.

**КЛЮЧОВІ СЛОВА:** метод, метрика, туберкульоз, Random Forest, Susceptible-Infectious-Recovered, моделювання, алгоритм.

## ЛІТЕРАТУРА

1. Duko B. The prevalence of depression among patients with tuberculosis: a systematic review and meta-analysis / B. Duko, A. Bedaso, G. Ayano // Ann Gen Psychiatry. – 2020. – Vol. 19(1). – P. 1. – Mode of access: https://doi.org/10.1371/journal.pone.0227472.
2. Association between tuberculosis and depression on negative outcomes of tuberculosis treatment: a systematic review and meta-analysis / [P. Ruiz-Grosso, R. Cachay, A. De La Flor et al.] // PLoS ONE. – 2020. – Vol. 15(1). – P. 1. – Mode of access: https://doi.org/10.1371/journal.pone.0227472.
3. The global prevalence of latent tuberculosis: a systematic review and meta-analysis / [A. Cohen, V.D. Mathiasen, T. Schön et al.]// Eur Respir J. – 2019. – Vol.54(3). – P. 1. – Mode of access: https://doi.org/10.1183/13993003.00655-2019.
4. Kiazyk S. Tuberculosis (TB): latent tuberculosis infection: an overview / S. Kiazyk, T. Ball // Canada Commun Dis Rep. – 2017. – Vol. 43(3–4). – P. 62. – Mode of access: https://doi.org/10.14745/ccdr.v43i34a01.
5. World Health Organization. Annual Report of Tuberculosis. Annual Global TB Report of WHO [Electronic resource]. – Access mode: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 (access date: 11.05.2024). – Title from the screen.
6. Whole genome sequencing-based analysis of tuberculosis (TB) in migrants: Rapid tools for crossborder surveillance and to distinguish between recent transmission in the host country and new importations / [E. Abascal, L. Pérez-Lago, M. Martínez-Lirola et al.] // Eurosurveillance. – 2018. – Vol. 24(4). – P. 1800005. – Mode of access: https://doi.org/10.2807/1560-7917.ES.2019.24.4.1800005.
7. Molecular surveillance of multi- and extensively drug-resistant tuberculosis transmission in the European Union from 2003 to 2011 / [J. L. De Beer, C. Ködmön, M. J. van der Werf et al.] // Eurosurveillance. – 2014. – Vol. 19(11). – P. 20742. – Mode of access: https://doi.org/10.2807/1560-7917.ES2014.19.11.20742.
8. Whole genome sequencing of multidrug-resistant Mycobacterium tuberculosis isolates collected in the Czech Republic, 2005–2020 / [M. Dohál, V. Dvořáková, M. Šperková et al.]// Sci Rep. – 2022. – Vol. 12(1). – P. 1–10. – Mode of access: https://doi.org/10.1038/s41598-022-11287-5.
9. Anti-tuberculosis drug resistance in Slovakia, 2018–2019: the first whole-genome epidemiological study / [M. Dohál, V. Dvořáková, M. Šperková et al.] // J Clin Tuberc Other Mycobact Dis. – 2022. – Vol. 26. – P. 100292. – Mode of access: https://doi.org/10.1016/j.jctube.2021.100292.
10. Alarming levels of multidrug-resistant tuberculosis in Ukraine: results from the first national survey / [E. Pavlenko, A. Barbova, A. Hovhannesyan et al.] // Int J Tuberc


DOI 10.15588/1607-3274-2024-4-4

Lung Dis. – 2018. – Vol. 22(2). – P. 197–205. – Mode of access: https://doi.org/ 10.5588/ijtld.17.0254.
11. Vyklyuk Ya. GeoCity – a New Dynamic-Spatial Model of Urban Ecosystem / Ya. Vyklyuk, D. Nevinskyi, N. Boyko // J. Geogr. Inst. Cvijic. – 2023. – Vol. 73(2). – P. 187–203. – Mode of access: https://doi.org/10.2298/IJGI2302187V.
12. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study / [M. Zignol, A. M. Cabibbe, A.S. Dean et al.] // Lancet Infect Dis. – 2018. – Vol. 18(6). – P. 675–683. – Mode of access: https://doi.org/10.1016/S1473-3099(18)30073-2.
13. Public health concerns about Tuberculosis caused by Russia/Ukraine conflict / [L. K. Paul, G. Nchasi, D. B. Bulimbe et al.] // Health Sci Rep. – 2023. – Vol.6(4). – P. 1. – Mode of access: https://doi.org/10.1002/hsr2.1218.
14. Tuberculosis in migrants moving from high-incidence to low-incidence countries: a population-based cohort study of 519,955 migrants screened before entry to England, Wales, and Northern Ireland. / [R. W. Aldridge, D. Zenner, P. J. White et al.] // Lancet. – 2016 – Vol. 388(10059). – P. 2510. – Mode of access: https://doi.org/10.1016/S0140-6736(16)31008-X.
15. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage / [M. Merker, C. Blin, S. Mona et al.] // Nat Genet. – 2015. – Vol. 47(3). – P. 242–249. – Mode of access: https://doi.org/10.1038/ng.3195.
16. Next-generation sequencing for characterizing drug resistance-conferring mycobacterium tuberculosis genes from clinical isolates in the Ukraine / [L. T. Daum, O. S. Konstantynovska, O. S. Solodiankin et al.] // J Clin Microbiol. – 2018. – Vol. 56(6). – P. 1. – Mode of access: https://doi.org/10.1128/JCM.00009-18.
17. Mycobacterium tuberculosis Beijing lineage and risk for tuberculosis in child household contacts. / [C. C. Huang, A. L. Chu, M. C. Becerra et al.] // Peru. Emerg Infect Dis. – 2020. – Vol. 26(3). – P. 568. – Mode of access: https://doi.org/10.3201/eid2603.191314.
18. Applying geospatial multi-agent system to model various aspects of tuberculosis transmission / [Y. Vyklyuk, I. Semianiv, D. Nevinskyi et al.] // New Microbes and New Infections. – 2024. – Vol. 59. – P. 101417. – Mode of access: https://doi.org/10.1016/j.nmni.2024.101417.
19. Standard genotyping overestimates transmission of mycobacterium tuberculosis among immigrants in a low-incidence country / [D. Stucki, M. Ballif, M. Egger et al.] // J Clin Microbiol. – 2016. – Vol.54(7). – P. 1862–70. – Mode of access: https://doi.org/10.1128/JCM.00126-16.
20. Jackson S. Effects of migration on tuberculosis epidemiological indicators in low and medium tuberculosis incidence countries: a systematic review / S. Jackson, Z. Kabir, C. Comiskey // J Clin Tuberc Other Mycobact Dis. – 2021. – Vol. 23. – P. 2405–5794. – Mode of access: https://doi.org/10.1016/j.jctube.2021.100225.


DOI 10.15588/1607-3274-2024-4-4

UDC 519.87:004.02

# ABOUT OF THE ANNEALING METHOD USING FOR THE TRAVELING SALESMAN PROBLEM SOLUTION WITH THE FUZZY TIME PERCEPTION

**Ivohin E. V.** – Dr. Sc., Professor, Professor of the Department of System Analysis and Decision Support Theory, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine.

**Adzhubey L. T.** – PhD, Associate Professor, Associate Professor of the Department of Computational Mathematics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine.

**Makhno M. F.** – PhD, Associate Professor, Ass. Professor of the Department of System Analysis and Decision Support Theory, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine.

**Rets V. O.** – Postgraduate student of the Department of System Analysis and Decision Support Theory, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine.

## ABSTRACT

**Context.** The article considers a technique for the use of fuzzy numbers and the annealing method for solving the traveling salesman problem, which is formulated as the problem of finding a route to visit a given number of cities without repetitions with a minimum duration of movement. The task of formalizing the algorithm for solving the traveling salesman problem by the annealing method using fuzzy numbers for subjective time perception is posed. The use of fuzzy numbers to increase the accuracy to represent real-world circumstances is proposed.

**Objective.** The goal of the work is to develop an algorithm for solving the traveling salesman problem based on the implementation of the annealing method with fuzzy numbers representing the subjective time perception for traveling between the cities with the minimum perceived duration of movement along the route.

**Method.** This paper proposes a method for solving the traveling salesman problem by the annealing method with fuzzy numbers for subjective time perception. A scheme for formalizing the procedure for solving the traveling salesman problem with the minimal perceived duration of movement along the route is described. A variant of the original traveling salesman problem is proposed, which consists in using fuzzy numbers to represent the uncertainty and subjective time perception in traveling between cities as opposed to regular crisp numbers to show regular distance and/or time of traveling. The results of the proposed algorithm for calculating solutions to the traveling salesman problem with minimization of the perceived duration of movement are presented, the obtained solutions are compared with the solutions found by other heuristic methods.

**Results.** The method for solving the traveling salesman problem using the annealing method with fuzzy numbers for subjective time perception is developed. A variant of the original traveling salesman problem is proposed, which consists in using fuzzy numbers to represent the uncertainty and subjective time perception in traveling between cities as opposed to regular crisp numbers to show regular distance and/or time of traveling. The application of fuzzy numbers makes it possible to perform calculation over possibly uncertain or subjective data, making results more accurate in the case of realistic deviations from the expected mean values in distance coverage. The results of the proposed algorithm for calculating solutions to the traveling salesman problem with minimization of the perceived duration of movement are presented, the obtained solutions are compared with the solutions found by other heuristic methods.

**Conclusions.** The paper considers a method for formalizing the algorithm for solving the traveling salesman problem using fuzzy numbers for subjective time perception. The use of fuzzy numbers to increase the accuracy to represent real-world circumstances is proposed. The scheme for formalizing the procedure for solving the traveling salesman problem with the minimal perceived duration of movement along the route is described. A variant of the original traveling salesman problem is proposed, which consists in using fuzzy numbers to represent the uncertainty and subjective time perception in traveling between cities as opposed to regular crisp numbers to show regular distance and/or time of traveling.

**KEYWORDS:** traveling salesman problem, fuzzy numbers, simulated annealing, combinatorial optimization, subjective perception of time, imprecision, uncertainty.

## ABBREVIATIONS

TSP is a traveling salesman problem.

## NOMENCLATURE

$p$ is a cyclic permutation of numbers;

$j_i$ is a city number;

$n$ is a number of cities;

$d_{ij}$ is a travel time between all pairs of vertices;

$D$ is a matrix of moving cost (distances or times);

$i, j$ are the indexes;

$I$ is a set of vertex indices;

$X$ is a binary matrix of transitions between vertices;

$x_{ij}$ are the elements of matrix X, which equal to 0 or 1;

$v_i$ is a vertex of graph, $i = \overline{1,n}$;

$t$ is a moment of time;

$S$ is a set of all system state;

$f(s)$ is a state change function;

$s_i$ is a system state on $i$-th step;

$s_k$ is a new state (candidate);

$t_{\min}$ is minimal temperature;

$t_{\max}$ is an output temperature;


DOI 10.15588/1607-3274-2024-4-5

$t_i$ is a current temperature of annealing process;
$T(t)$ is a temperature change function;
$E(s)$ is an objective function value;
$\tilde{A}$ is a fuzzy set;
$E$ is a set of numbers;
$\mu_{\tilde{A}}(x)$ is a membership function;
$(a, b, c)$ is a triangle fuzzy number;
$F1(\tilde{A})$ is a rank of a fuzzy number $\tilde{A}$;
$g(x)$ is a weight function;
$M$ is a random number.

## INTRODUCTION

The way decisions are made in society in many cases depends on the emotional state of a person. Feelings are like a reference point that is determined by a goal that is influenced by various factors. Emotions can be the reason for behavior that is appropriate for a particular situation, even when it is not the most efficient, but allows you to avoid any consequences that may arise from exceeding a certain time limit.

Special attention should be paid to these factors in the processes of formation and improvement of many theoretical ideas in the field of modeling human behavior, one of which is the adaptation of physical and mathematical models to real life. This makes it possible to combine the power of computational methods with the peculiarities of human behavior. Such tasks are common in the context of the application of artificial intelligence methods and algorithms, the creation of decision-making support systems, the resolution of resource allocation issues taking into account the human factor, etc.

Time is an important resource in activities involving human participation. Estimation of time intervals is fundamental to understanding time frames, even though the exact boundaries of the interval may not be defined until the process reaches a certain stage. Thus, a period of time is usually defined by an indefinite interval that can be roughly predicted given the nature of the passage of time, if the given limits of the interval are taken as a given. To measure time intervals, they can be expressed in phrases such as “quick response”, “normal timing” or “long wait”. This means that when solving problems that require verbal terms to refer to time, it is important to take time variation into account. It is clear that emotions have a great influence on the understanding of time in processes that involve a person [1].

In order to find the most successful or effective solution to problems, it is necessary to take into account factors that affect human emotions and, therefore, the speed of time perception, resource allocation and calendar planning. The paper proposes to develop an approach to the formalization of accounting for the flow of time based on fuzzy numbers and to apply it to solving certain fuzzy optimization problems related to taking into account the fuzzy perception of time arising from the subjectivity and irregularity of the time count.

**The object of study** is the process of optimal route search for the fuzzy traveling salesman problem with a minimum duration of movement.

**The subject of study** is the algorithm for solving the fuzzy traveling salesman problem based on the combinatorial methods in combination with triangle fuzzy numbers.

**The purpose of the work** is to develop an algorithm for solving the fuzzy traveling salesman problem using one of the combinatorial methods of the approximate solution of the problem in combination with fuzzy numbers as a way to define subjective perception of time to travel between cities.

## 1 PROBLEM STATEMENT

The traveling salesman problem is one of the most famous computational optimization problems. The task is to find the shortest route that passes through each city exactly once for a given number of cities. The search for such a path was formulated as a mathematical problem in 1930 and is still one of the most intensively researched optimization problems [2].

The number of alternative paths for a TSP with $n$ nodes, where the nodes are cities and the edges are the cost of moving between two cities, is $(n–1)!$. Therefore, even for small problems, such as the one presented with only 20 nodes, the number of alternative paths is about $1.2*10^{17}$, to which there is no adequate computing power to explore through exhaustive enumeration.

The traveling salesman problem is one of the famous combinatorial problem [3]. To reduce the problem to a general form, we number the cities by numbers (1, 2, 3, ..., $n$), and describe the traveling salesman’s route by a cyclic permutation of numbers $p = (j_1, j_2, ..., j_n, j_1)$, where all $j_1, ..., j_n$ are different numbers.

The set of cities can be considered as the vertices of some graph with given distances (or travel time) between all pairs of vertices $d_{ij}$ that form the matrix $D=(d_{ij})$, $i,j=\overline{1,n}$. We assume that the matrix is symmetric. The formal problem then is to find the shortest route (in time or length) $t$ that goes through each city and ends at the starting point. In this formulation, the problem is called the closed traveling salesman problem, which is a well-known mathematical integer programming problem.

Let us formulate a mathematical model of the TSP problem. Let $I=\{1,...,n\}$ be the set of vertex indices of the problem graph. The objective function is the total distance or time of the route, including all the vertices of the task graph. The parameters of the problem are the elements of the matrix $D=(d_{ij})$, $i,j \in I$.

Shift tasks are elements of the binary matrix of transitions between vertices $X = \{x_{ij}\}$, $i, j \in I$, which are equal to 1 if there is an edge $(v_i, v_j)$ in the constructed route for the task, 0 otherwise [4]. The shortest route in terms of distance or time is optimal:


DOI 10.15588/1607-3274-2024-4-5

$$\sum_{i=1}^{n}\sum_{j=1,\,j\neq i}^{n} d_{ij}x_{ij} \to \min \qquad (1)$$

with constraints

$$\sum_{j=1,\,j\neq i}^{n} x_{ij} = 1,\ i=\overline{1,n},$$

$$\sum_{i=1,\,i\neq j}^{n} x_{ij} = 1,\ j=\overline{1,n}, \qquad (2)$$

$$v_i - v_j + nx_{ij} \le n-1,\ 1 \le i \ne j \le n.$$

The last inequality ensures the connectivity of the vertex traversal route; it cannot consist of two or more unconnected parts.

The dynamic traveling salesman problem (DTSP) is a TSP defined by a dynamic cost (distance) matrix as follows:

$$D = \left\{d_{ij}(t)\right\}_{n(t)\times n(t)},$$

where $d_{ij}(t)$ is the amount of moving cost from city (node) $i$ to city $j$ at time $t$. In this definition, the number of cities $n(t)$ and the cost matrix are time dependent. The traveling salesman's dynamic problem is to find a minimum-cost route that contains all $n(t)$ nodes.

In other words, having all $n(t)$ nodes $j_1,\ldots,\ j_{n(t)}$ and the corresponding cost matrix $D(t)=(d_{ij}(t))$, $i,j=\overline{1,n(t)}$, we need to find a route with the minimum cost containing all $n(t)$ points, where $t$ is the moment of time, $d_{ij}(t)$ is the distance or time between the points $i$ and $j$:

$$\sum_{i=1}^{n(t)}\sum_{j=1,\,j\neq i}^{n(t)} d_{ij}(t)x_{ij} \to \min \qquad (3)$$

with above constraints (2).

The change in the cost matrix $D$ over time is a continuous process. Practically, in order to build analytical models, it is necessary to discretize this process of changes. Thus, $D$ becomes a series of optimization problems $D(t_k)=(d_{ij}(t_k))$, $i,j=\overline{1,n(t_k)}$, $k$ = 0, 1, 2, …, $m$–1, with time windows $[t_k, t_{k+1}]$, where $\{t_k\}_{i=0}^{m}$ is a sequence of time points.

## 2 REVIEW OF LITERATURE

Algorithms that allow solving the problem of finding the optimal route are divided into exact and heuristic. In the case of exact methods, the search for solutions is based on optimization methods such as linear programming, dynamic programming, or the branch and bound method [5]. However, it is convenient to use exact methods only for small-scale problems (for example, for the purpose of primary design of a small-sized transport network), since their implementation requires large computing power.

Heuristic methods do not guarantee finding an optimal solution, but are aimed at quickly finding a locally optimal solution. Traditionally, "trial and error" approaches, such as random search or greedy algorithm, are used to quickly explore the solution space and find a promising solution [6]. Heuristics are more flexible and can be applied to larger problems, but the solution they offer may not be optimal. Among such heuristic methods, attention should also be paid to methods that imitate biological (ant colony algorithm and genetic algorithm [7, 8]) or physical processes [9, 10].

One of the methods of solving the traveling salesman problem using the combinatorial optimization technique is the annealing method [9]. By analogy with the annealing process of various physical materials, in which by raising its temperature to a high level and then gradually lowering it, the algorithm accidentally disturbs the output path ("heating") for further gradual lowering of the "temperature" [10].

When modeling the annealing process, the analog of temperature is the level of randomness, with the help of which changes are made to the path, which in the future improves in its duration. When the "temperature" of the process is high, changes occur to avoid the danger of reaching a local minimum, followed by control at the optimal value as the "temperature" is successively reduced. "Temperature" decays in a series of steps on an exponential decay curve, with each step the temperature being lower than before.

## 3 MATERIALS AND METHODS

Let's describe an annealing method. The approach implemented in the simulated annealing method is borrowed from physical processes. It is based on the process of crystallization of a substance, which metallurgists found to increase the homogeneity of the metal.

As is known, metals have crystal lattices that determine the geometric position of the atoms of the substance. The set of positions of all atoms will be called the state of the system; each state corresponds to a certain energy level. The purpose of annealing is to bring the system to the state with the lowest energy. The lower the energy level, the "better" the crystal lattice, that is, the fewer defects it has and the stronger the metal.

During annealing, the metal is first heated to any temperature, which causes the atoms of the crystal lattice to leave their positions. Slow and controlled cooling then begins. Atoms tend to get into a state with lower energy, but with a certain probability they can go into a state with higher energy. This probability decreases with temperature. The transition to a worse state, oddly enough, helps as a result to find a state with less energy than the original one. The process ends when the temperature drops to the set value.

Such a complex scheme with probabilities of transition from point to point is necessary so that the algorithm


DOI 10.15588/1607-3274-2024-4-5

does not get stuck on a local minimum, taking it for a global optimum. To get out of this situation, you need to increase the energy of the system from time to time. At the same time, the general tendency to search for the lowest energy remains. This is the essence of the simulated annealing method.

To describe the algorithm scheme for formalizing the simulated annealing method, we introduce the notation:

$S$ – set of all system state;

$f(s)$ – state change function;

$s_i$ – system state on $i$-th step;

$s_k$ – new state (candidate);

$t_{\min}$, $t_i$, $t_{\max}$ – minimal, current and output temperature respectively;

$T(t)$ – temperature change function;

$E(s)$ – objective function value.

The algorithm starts working from the initial state $s_1$, with the initial temperature $t_1 = t_{\max}$ and with the specified minimum temperature $t_{\min}$.

For every steps with numbers $i$=1,2,… while $t_i > t_{\min}$ repeat:

1) $s_k = f(s_{i-1})$;

2) $\Delta E = E(s_k) - E(s_{i-1})$;

3) if $\Delta E < 0$, then $s_i = s_k$;

4) otherwise, acceptance of a new state occurs with some probability $\exp(-\Delta E / t_i)$;

5) choose a random number $M$ on interval (0,1);

6) if $\exp(-\Delta E / t_i) > M$, perform the transition $s_i = s_k$, otherwise, go to the next step;

7) reduce the temperature $t$: $t_{i+1} = T(t_i)$;

8) return the last state $s_i$, $i = i+1$.

**Fuzzy passage of time**

In everyday life, expressions such as “almost six”, “quite tall”, “not short enough” are often used to define a certain size in an approximate format. As a result, this method of evaluation requires the formalization of insufficiently clearly defined evaluations for their practical application in mathematical models. For this purpose, you can use concepts that allow you to present the subjective or intuitive meaning of fuzzy concepts in a constructive way. One of these concepts of uncertainty formalization is fuzzy numbers [11].

Fuzzy numbers are used to obtain results in problems related to decision-making and analysis. Fuzzy numbers defined in the number space are an extension of real numbers and have their own properties that can be attributed to number theory. To understand fuzzy numbers and their subspecies – triangular and parabolic numbers, consider the concept of a fuzzy set.

Let $E$ be a set with a finite or infinite number of elements. Let $A$ be the set contained in $E$. Then the set of ordered pairs $(x, \mu_{\tilde{A}}(x))$ defines a fuzzy subset $\tilde{A}$ for $E$, where $x$ – is a member of $E$, and $\mu_{\tilde{A}}(x)$ – degree of belonging of $x$ to $A$. The set of elements from $A$ for which $\mu_{\tilde{A}}(x)>0$ form the support of a fuzzy set.

A fuzzy number is a generalization of an ordinary real number. It refers to a connected set of possible values, where each possible value has its own weight between 0 and 1. Thus, a fuzzy number is a special case of a convex normalized fuzzy set in the space of real numbers. Among the possible types of fuzzy numbers, triangular and parabolic numbers are considered in the work.

A fuzzy number $\tilde{A} = (a, b, c)$ is called a triangular fuzzy number if its membership function looks like this:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, x < a; \\ (x-a)/(b-a), a \le x \le b; \\ (c-x)/(c-b), b \le x \le c; \\ 0, x > c. \end{cases} \quad (4)$$

Above the triangular numbers (Fig. 1), you can determine the main arithmetic operations for further use in calculations.

Let $\tilde{A} = (a, b, c)$ and $\tilde{B} = (a1, b1, c1)$ be two triangular numbers. Then:

– The sum is defined as $\tilde{A} + \tilde{B} = (a+a1, b+b1, c+c1)$.

– The difference is defined as $\tilde{A} - \tilde{B} = A + (-\tilde{B}) = (a-c1, b-b1, c-a1)$, where $-\tilde{B} = (-c1, -b1, -a1)$ is defined as the opposite of $B$.

In other words, opposite triangular numbers and their sum and difference are also triangular numbers. It is also worth noting that the results of inversion and multiplication of triangular numbers do not preserve this property and do not always represent triangular numbers.

The parabolic number (Fig. 1) is given similarly and has the same properties, but has a different membership function:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, x < a, \\ -((x-b)/(a-b))^2 + 1, a \le x \le b, \\ -((x-b)/(c-b))^2 + 1, b \le x \le c, \\ 0, x > c. \end{cases} \quad (5)$$

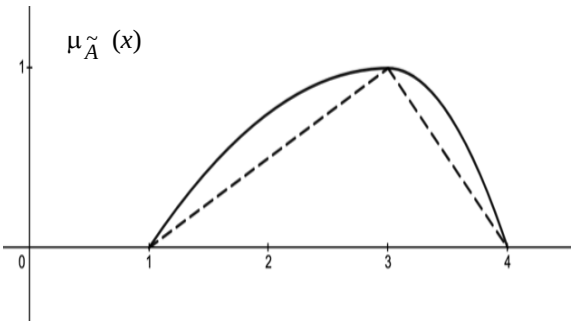



Figure 1 – Example of triangle and parabolic fuzzy number


DOI 10.15588/1607-3274-2024-4-5

Although uncertainty information can be formalized using fuzzy numbers, the decision-making procedure must be precise. For example, the final output of fuzzy systems and the selection of appropriate solutions should be justified on the basis of the value characterized by the confidence (importance) indicator. To obtain a clear value, methods of calculating ranks (defuzzification) of fuzzy numbers are used, which are essentially "crisp" representative numbers, and can be used as generalized values for further calculations. One of the methods for calculating the rank of a fuzzy number is the Jaeger method, which calculates the Jaeger rank of the first type in the form [12]:

$$F1(\tilde{A}) = \frac{\int_0^1 g(x)\mu_{\tilde{A}}(x)dx}{\int_0^1 \mu_{\tilde{A}}(x)dx}, \quad (6)$$

where $g(x)$ is a weight function that measures the importance of the value of $x$. If $g(x) = x$, the index can be considered as the geometric center $\tilde{A}$, as shown in Figure 1. The support of the fuzzy number in this case is the segment [0, 1]. If the reference sets of the fuzzy numbers being compared do not coincide with [0, 1], then they can be scaled by dividing each of the numbers by max[sup $S_{\tilde{A}_i}$ ], where $S_{\tilde{A}_i}$ denotes the $i$-th reference set fuzzy number. Using this scaling procedure will give a factor of 1/ max[sup $S_{\tilde{A}_i}$ ] (1, if no scaling is used). The limits of integration in this case will be min[inf $S_{\tilde{A}_i}$ ] and max[sup $S_{\tilde{A}_i}$ ], respectively.

If $g(x) = x$ and the fuzzy number is triangular, the index $F1$ reduces to a simpler form:

$$F1(\tilde{A}) = \frac{1}{3}(a+b+c), \quad (7)$$

and in the case of a parabolic number:

$$F1(\tilde{A}) = \frac{1}{8}(3a+2b+3c), \quad (8)$$

where $a = inf\ S_{\tilde{A}_i}$ , $\mu_{\tilde{A}}(b) = 1$ , $c$ = sup $S_{\tilde{A}_i}$ .

This method of defuzzification is also called the method of the center of gravity (COG) [13] (Fig. 2). Among other well-known methods, it is also worth noting the bisector of area (BOA) method [14], according to which there is such a value of $x$ that a vertical line drawn through it divides the fuzzy number into two equal parts by their area.

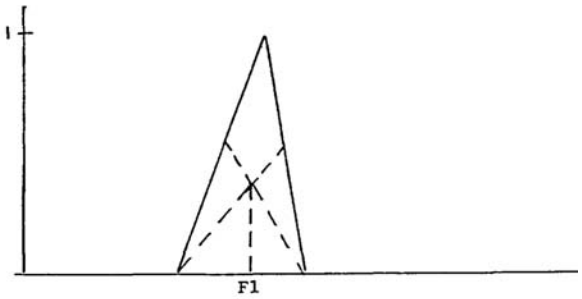



Figure 2 – Defuzzification by method of the gravity center

## 4 EXPERIMENTS

In the course of the study, the described algorithm was implemented based on the annealing method using fuzzy numbers to represent the subjective perception of the passage of time on road sections between cities. A multithreaded Python implementation is proposed for numerical calculations. In the process of work, the method of calculating the rank of fuzzy numbers is chosen and the rank values of different methods (use of the peak abscissa, BOA, COG) are compared with the average value of random route passes. It was concluded that the best result was obtained by the center of gravity (COG) method. A comparison of the route estimation methods is given in the table, where the random route characterizes the time of travel along the route taking into account the average speed, the calculated route is the duration estimate obtained by the chosen method (see Table 1).

Table 1 – Rezults of the route estimation

| Method | Random route | Calculated route |
|---|---|---|
| The peak abscissa | 5367.78 | 5046.0 |
| BOA (triangular FN) | 5332.54 | 5291.72 |
| BOA (parabolic FN) | 5369.91 | 5341.70 |
| COG (triangular FN) | 5332.40 | **5332.86** |
| COG (parabolic. FN) | 5369.67 | **5368.72** |

During the program processing, three possible approaches to finding solutions are compared (using crisp, triangular, and parabolic numbers, respectively). As initial conditions, the TSPLib library was used, which has known TSP conditions in its catalog in the form of arrays of coordinates or matrices of the conditional distances between cities (e.g. u16, fr4, etc.). Fuzzy initial conditions were randomly generated with possible deviation from the expected value in either direction. To test the proposed approach, the time of the best constructed results for each type of fuzzy numbers was compared with the average value of the time taken for 10^5 random passes along the constructed route.

## 5 RESULTS

The results of the numerical experiments are shown in the Table 2, in which the best solutions are defined with actual time for trip for the different conditional distances between cities.


DOI 10.15588/1607-3274-2024-4-5

Table 2 – Rezults of the method's comparison

| Task | Fuzzy number type | Estimated time | Actual time |
|---|---|---|---|
| u16 | crisp | 6859.0 | 7247.38 |
| | triangular | 6859.0 | 7129.38 |
| | **parabolic** | 6859.0 | **7128.92** |
| fr4 | crisp | 5046.0 | 5369.57 |
| | triangular | 5071.0 | 5369.30 |
| | **parabolic** | 5070.0 | **5363.88** |
| pr76 | crisp | 108273.0 | 115105.75 |
| | **triangular** | 108894.0 | **114102.55** |
| | parabolic | 109295.0 | 114563.41 |
| rd100 | crisp | 8185.0 | 8653.50 |
| | **triangular** | 7975.0 | **8390.85** |
| | parabolic | 8049.0 | 8447.22 |
| rd400 | crisp | 18070.0 | 19087.33 |
| | **triangular** | 18089.0 | 19008.33 |
| | **parabolic** | 17808.0 | **18713.72** |

Thus, it was concluded that the use of fuzzy numbers in the annealing algorithm allows to obtain constructive results when solving the traveling salesman problem with fuzzy input parameters.

## 6 DISCUSSION

Given that the "cost" of travel between cities in time measurement can vary depending on the situation, a more accurate representation of such cost can be given in the form of triangular or parabolic numbers. If the subjective perception of time is chosen as the value, the relative duration of the trip between cities may vary depending on the factors affecting the path – traffic jams, bad weather, etc. Note that even in a simpler perception of the dynamic duration of the road between cities, when the actual time required to cover the path at the recommended average speed is measured, the same factors change the given duration, and therefore it makes sense to represent the studied travel time in the form of triangular or parabolic numbers.

Using one of the combinatorial methods of the approximate solution of the traveling salesman problem in combination with fuzzy numbers (and the corresponding method of calculating their rank), it is possible to achieve an effective result from the construction of the optimal path taking into account the dynamic features of roads between destinations. At the same time, better calculations can be obtained when using fuzzy parabolic numbers, since their essence is closer to reality. For the subjective overestimation or underestimation of the perception of the passage of time, the rule is valid: the greater the possible deviation in perception, the less likely it is to be obtained. For the numerical implementation of the actions of the annealing algorithm on fuzzy numbers that determine the time perception of the duration of movement between cities, operations according to the above schemes are used, and the different routes formed at the same time are compared with each other by finding and comparing the ranks of fuzzy numbers using one of the specified methods.

## CONCLUSIONS

This paper investigates the use of fuzzy numbers and the annealing method to find a solution to the traveling salesman problem, which involves finding the shortest route for a given set of cities. Fuzzy numbers are used to model the inaccuracy and uncertainty of input data, and an annealing method is proposed to find solutions. The solutions obtained on the basis of the developed program in the Python language were analyzed. A comparison of the results of the TSP problem using crisp and fuzzy numbers using the annealing method was carried out. The results of numerical experiments are given, which show that the use of fuzzy numbers, in particular triangular and parabolic, with the annealing method leads to a significant improvement in the results of the TSP problem compared to the use of crisp numbers. This approach can be applied to real-world optimization problems involving imprecise or uncertain data and can be useful for optimizing processes with subjective time perception. A conclusion was made about the need for further research using the theory of fuzzy numbers, in particular in the direction of the correct choice of the type of numbers in accordance with the conditions of the task. Another direction of research involves further development of the proposed methodology for solving fuzzy dynamic traveling salesman problems and the use of other effictive (for example, genetic), approximate and greedy methods.

## ACKNOWLEDGMENTS

The work is the result of research carried out by the Faculty of Computer Science and Cybernetics of the Taras Shevchenko National University of Kyiv, the Department of Information Systems and Technologies of the National Transport University and the travel company Bon Voyage during the implementation of the project for building optimal multi-day tourist routes.

## REFERENCES

1. Schirmer A. How emotions change time, *Frontiers in Integrative Neuroscience*, 2011, № 5, P. 58. DOI: https://doi.org/10.3389/fnint.2011.00058
2. Schrijver A. Theory of Linear and Integer Programming. Wiley, 1998, 484 p. DOI: 10.2307/253980
3. Matai R. Singh S. P., Mittal M. L. Traveling salesman problem: an overview of applications, formulations, and solution approaches, *Traveling salesman problem, theory and applications*, 2010, № 1, pp. 11–17. DOI:10.5772/12909
4. Vanderbei R. J. Linear programming: Foundations and extensions. Springer, 2014, 414 p. DOI: 10.1007/978-3-030-39415-8
5. Korte B., Vygen J. Combinatorial Optimization: Theory and Algorithms (Algorithms and Combinatorics). Berlin, Heidelberg, Springer, 2018, 455 p. DOI: 10.1007/978-3-662-56039-6
6. Moss L. T., Atre Shaku Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applica-


DOI 10.15588/1607-3274-2024-4-5

tions. Addison-Wesley Professional, 2003, 576 p. ISBN: 978-0-201-78420-6
7. Buontempo F. Genetic Algorithms and Machine Learning for Programmers. Pragmatic Bookshelf, 2019, 236 p. ISBN: 9781680506204
8. Kshemkalyani A. D., Singhal Mukesh Distributed Computing: Principles, Algorithms, and Systems. Cambridge University Press, 2008, 756 p. DOI: 10.1017/CBO9780511805318
9. Rai S., Ettam R.K.Simulation-based optimization using simulated annealing for optimal equipment selection within print production environments, *2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World, December, 2013: proceedings*, 2013, pp. 1097–1108. DOI: 10.1109/WSC.2013.6721499
10. Grabusts P., Musatovs J., Golenkov V. The application of simulated annealing method for optimal route detection between objects, *Procedia Computer Science*, 2019,V. 149, pp. 95–101. DOI: https://doi.org/10.1016/j.procs.2019.01.112
11. Heilpern S. Representation and application of fuzzy numbers, *Fuzzy sets and Systems*, 1997, №91(2), pp. 259–268. DOI: https://doi.org/10.1016/S0165-0114(97)00146-2
12. Zadeh L.A. Fuzzy sets, *Information and Control*, 1965, №8. pp. 338–353. DOI: http://dx.doi.org/10.1016/S0019-9958(65)90241-X
13. Tofigh A., Saneifard R. Defuzzification method for ranking fuzzy numbers based on center of gravity, *Iranian Journal of Fuzzy Systems*, 2012, №9(6), pp. 57–67. DOI: 10.22111/IJFS.2012.113
14. Nejad A. M., Mashinchi M. Ranking fuzzy numbers based on the areas on the left and the right sides of fuzzy number, *Computers & Mathematics with Applications*, 2011, Vol. 61, №2, pp. 431–442. DOI:10.1016/j.camwa.2010.11.020

Received 05.09.2024.
Accepted 01.11.2024.

УДК 519.87:004.02

## ПРО ВИКОРИСТАННЯ МЕТОДУ ВІДПАЛУ ДЛЯ РОЗВ'ЯЗАННЯ ЗАДАЧІ КОМІВОЯЖЕРА З НЕЧІТКИМ СПРИЙНЯТТЯМ ЧАСУ

**Івохін Є. В.** – д-р фіз.-мат. наук, професор, професор кафедри системного аналізу та теорії прийняття рішень Київського національного університету імені Тараса Шевченка, Київ, Україна.

**Аджубей Л. Т.** – канд. фіз.-мат. наук, доцент, доцент кафедри обчислювальної математики Київського національного університету імені Тараса Шевченка, Київ, Україна.

**Махно Л. Т.** – канд. техн. наук, доцент, доцент кафедри системного аналізу та теорії прийняття рішень Київського національного університету імені Тараса Шевченка, Київ, Україна.

**Рець В. О.** – аспірант кафедри системного аналізу та теорії прийняття рішень Київського національного університету імені Тараса Шевченка, Київ, Україна.

### АНОТАЦІЯ

**Актуальність.** Інтеграція нечітких чисел в алгоритми має вирішальне значення для вдосконалення обчислювальних методологій. Нечіткі числа з властивою їм неточністю пропонують більш реалістичне уявлення про явища реального світу. Адаптація та інноваційні алгоритми для включення нечітких чисел є важливими для вирішення складних проблем, коли дані можуть бути неточними або неоднозначними. Це вдосконалення допомагає більш обґрунтовано приймати рішення зважаючи на тонкощі реального світу, що у свою чергу сприяє прогресу в різних сферах і дозволяє проводити дослідження у контексті суб'єктивного сприйняття часу.

**Ціль.** Мета роботи – розробити алгоритм розв'язання задачі комівояжера з використанням нечітких чисел для формалізації невизначеності та неточності вхідних даних, пов'язаної з впливом суб'єктивності в оцінках тривалості необхідних проміжків часу.

**Метод.** У статті розглянуто метод відпалу з нечітким представленням часу для розв'язання нечіткої задачі комівояжера, що формулюється як задача знаходження маршруту відвідування заданої кількості міст без повторень з мінімальною тривалістю руху з нечіткими числами, що представляють час, необхідний для подолання відстаней між містами. Поставлено та вирішено задачу формалізації алгоритму розв'язання проблеми комівояжера на основі методу відпалу з використанням нечітких чисел. Запропоновано можливі методи апроксимації нечітких чисел в контексті поставленої задачі. Розроблено конструктивний алгоритм розв'язання задачі. Проведено обчислювальні експерименти.

**Результати.** Розроблено метод розв'язання задачі комівояжера з використанням методу відпалу та нечітких чисел. Запропоновано використання нечітких чисел для формалізації невизначеності та неточності вхідних даних, пов'язаної з впливом суб'єктивності в оцінках тривалості необхідних проміжків часу. Представлено результати розрахунків за допомогою запропонованого алгоритму в задачах комівояжера з мінімізацією суб'єктивної тривалості руху, показано можливі методи апроксимації нечітких чисел та їх порівняння в контексті поставленої задачі, проведено порівняння отриманих розв'язків із розв'язками, знайденими за допомогою інших евристичних методів.

**Висновки.** У статті розглянуто метод формалізації алгоритму розв'язання задачі комівояжера з використанням алгоритму методу відпалу та нечітких чисел. Запропоновано використання нечітких чисел для формалізації невизначеності та неточності вхідних даних, пов'язаної із впливом суб'єктивності в оцінках тривалості необхідних проміжків часу. Описано схему формалізації процедури використання методу відпалу з нечіткими числами, що представляють суб'єктивне представлення часу, необхідного для подолання відстаней між містами.

**КЛЮЧОВІ СЛОВА:** задача комівояжера, нечіткі числа, метод відпалу, комбінаторна оптимізація, суб'єктивне сприйняття плину часу, неточність, невизначеність.


DOI 10.15588/1607-3274-2024-4-5

**ЛІТЕРАТУРА**

1. Schirmer A. How emotions change time/ A. Schirmer // Frontiers in Integrative Neuroscience. – 2011. – № 5. – P. 58. DOI: https://doi.org/10.3389/fnint.2011.00058
2. Schrijver A. Theory of Linear and Integer Programming/ A. Schrijver. – Wiley, 1998. – 484 p. DOI: 10.2307/253980
3. Matai R. Traveling salesman problem: an overview of applications, formulations, and solution approaches / R. Matai, S. P Singh, M. L. Mittal // Traveling salesman problem, theory and applications. – 2010. – № 1. – P. 11–17. DOI:10.5772/12909
4. Vanderbei R. J. Linear programming: Foundations and extensions / R. J. Vanderbei. – Springer, 2014. – 414 p. DOI: 10.1007/978-3-030-39415-8
5. Korte B. Combinatorial Optimization: Theory and Algorithms (Algorithms and Combinatorics)/ B. Korte, J. Vygen. – Berlin, Heidelberg : Springer, 2018. – 455 p. DOI: 10.1007/978-3-662-56039-6
6. Moss L. T. Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications/ L. T. Moss, Shaku Atre. – Addison-Wesley Professional, 2003. – 576 p. ISBN: 978-0-201-78420-6
7. Buontempo F. Genetic Algorithms and Machine Learning for Programmers/ F. Buontempo. – Pragmatic Bookshelf, 2019. – 236 p. ISBN: 9781680506204
8. Kshemkalyani A. D. Distributed Computing: Principles, Algorithms, and Systems / A. D. Kshemkalyani, Mukesh Singhal. – Cambridge University Press, 2008. – 756 p. DOI: 10.1017/CBO9780511805318
9. Rai S. Simulation-based optimization using simulated annealing for optimal equipment selection within print production environments/ S. Rai, R. K. Ettam// 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World, December, 2013: proceedings. – 2013. – P. 1097–1108. DOI: 10.1109/WSC.2013.6721499
10. Grabusts P. The application of simulated annealing method for optimal route detection between objects / P. Grabusts, J. Musatovs, V. Golenkov// Procedia Computer Science. – 2019. – V. 149. – P. 95–101. DOI: https://doi.org/10.1016/j.procs.2019.01.112
11. Heilpern S. Representation and application of fuzzy numbers/ S. Heilpern // Fuzzy sets and Systems. – 1997. – №91(2). – P. 259–268. DOI: https://doi.org/10.1016/S0165-0114(97)00146-2
12. Zadeh L. A. Fuzzy sets / L. A. Zadeh // Information and Control. – 1965. – № 8. – P. 338–353. DOI: http://dx.doi.org/10.1016/S0019-9958(65)90241-X
13. Tofigh A. Defuzzification method for ranking fuzzy numbers based on center of gravity / A. Tofigh, R. Saneifard// Iranian Journal of Fuzzy Systems. – 2012. – №9(6). – P. 57–67. DOI: 10.22111/IJFS.2012.113
14. Nejad A. M. Ranking fuzzy numbers based on the areas on the left and the right sides of fuzzy number / A. M. Nejad, M. Mashinchi // Computers & Mathematics with Applications. – 2011. – Vol. 61, № 2. – P. 431–442. DOI:10.1016/j.camwa.2010.11.020


DOI 10.15588/1607-3274-2024-4-5

UDC 519.852.6

# INNOVATIVE IMPROVED APPROXIMATE SOLUTION METHOD FOR THE INTEGER KNAPSACK PROBLEM, ERROR COMPRESSION AND COMPUTATIONAL EXPERIMENTS

**Mamedov K. Sh.** – Dr. Sc., Professsor of Baku State University and Head of the Departament of the Institute of Control Systems Ministry of Science and Education of Azerbaijan.

**Niyazova R. R.** – Doctorant and Scientist of the Institute of Control Systems Ministry of Science and Education of Azerbaijan.

## ABSTRACT

**Context**. Mathematical models of many optimization problems encountered in economics and engineering are taken in the form of an integer knapsack problem. Since this problem belongs to the class of "NP-complete", that is, "hard to solve" problems, the number of operations required by known methods to find its optimal solution is exponential. This does not allow solving large-scale problems in real time. Therefore, various and fast working approximate solution methods of this problem have been developed. However, it is known that the approximate solution provided by those methods can differ significantly from the optimal solution in most cases. Therefore, after taking any approximate solution as a starting point, there is a demand to develop methods for its further improvement. Development of such methods has both theoretical and great practical importance.

**Objective.** The main purpose solving of this issue is as follows**.** The main purpose in performing this work is to first find an initial approximate solution of the problem using any known method, and then work out an algorithm for successively further improvement of this solution. For this purpose, the set of numbers with which the coordinates of the optimal solution and the found approximate solution can differ should be determined. After that, new solutions should be constructed by assigning possible values to the unknowns corresponding to the numbers in that set, and the best among these solutions should be selected. However, the algorithm for constructing such a solution should be simple, require a small number of operations, not cause difficulties from the point of view of programming, be new and be applicable to practical issues.

**Method.** The essence of the proposed method consists of the following. First, the initial approximate solution of the considered problem and the value of the objective function corresponding to this solution are found by a known rule. After that, the optimal solution of the problem is easily found by a known method, without taking into account the condition that the unknowns are integers**.** Obviously, this solution can take at most one coordinate fractional value. It is assumed that the coordinates of the optimal solution of the integer knapsack problem and the initial approximate solution may differ around a certain fractional coordinate of the optimal solution of the continuous problem. Then, the minimum number of non-zero coordinates and zero coordinates in the optimal solution is found. Corresponding theorems have been proved for this. It is assumed that the different coordinates of the optimal solution and the initial approximate solution located between those minimal numbers. Therefore, the best solution can be selected by successively changing the coordinates between those minimum numbers one by one.

**Results.** Extensive calculation experiments were conducted with the application of the proposed method.To have a high quality of this method was confirmed once again through experiments.

**Conclusions.** The proposed method is new, simple in nature, easy to consider from the programming point of view, and has important practical importance. Thus, we call this solution the innovative improved approximate solution.

**KEYWORDS:** Integer knapsack problem, initial approximation solution, minimum number of zeros and non-zero coordinates in optimal solution, innovative improvement of initial approximation solution, error estimation and experiments.

## NOMENCLATURE

$N$ – number of issues resolved;

$n$ – number of unknowns;

$a_j, c_j, d_j (j = \overline{1,n})$ and $b$ are given positive integers;

$x_j, (j = \overline{1,n})$ – $j$-th unknown;

$X^t$ – approximate solution;

$f^t$ – value of the approximate solution to the objective function;

$X^*$ – optimal solution of the problem;

$f^*$ – optimal value of the objective function;

$X^{lp}$ – optimal solution of the appropriate linear programming problem;

$k$ – number of the fractional coordinate in the numerical $x^{lp}$ solution;

$f^{lp}$ – optimal value of the objective function in the appropriate linear programming problem;

$x_k$ – coordinate , that fractional coordinate in the $x^{lp}$ solution;

$p$, $q$, $\delta$ – mentioned positive integer;

$\delta_k^1, \delta_k^0$ – certain integers expressed as a percentage;

$n(1)$ and $n(0)$ – minimal number of non-zero coordinates and zeros, respectively, in the optimal solution of the considered problem;

$\omega_1, \omega_2$ – certain set of numbers, respectively;

$n(\omega_1), n(\omega_2)$ – number of elements of matching sets;

$\underline{X^*}$ and $\underline{X^{lp}}$ – optimal solutions of certain problems;


DOI 10.15588/1607-3274-2024-4-6

$n_1$ and $\bar{n}$ – respectively, is the minimum number of non-zero coordinates and the maximum number of zero coordinates in the optimal solution of a certain problem;

$X^{it}$ – innovative improved final solution;

$f_i^t$ – value of the objective function according to the solution $X^{it}$;

$\delta$ and $\delta^i$ respectively, are the relative errors (in percent) of the initial approximate solution and the innovative improved solution from the optimal solution.

## INTRODUCTION

Consider the following integer knapsack problem:

$$\sum_{j=1}^{n} c_j x_j \to \max, \tag{1}$$

$$\sum_{j=1}^{n} a_j x_j \le b, \tag{2}$$

$$0 \le x_j \le d_j, \quad (j=\overline{1,n}) \tag{3}$$

$$x_j - \ integer \ (j=\overline{1,n}). \tag{4}$$

Here, without breaking generality, we accept that – $a_j > 0, c_j > 0, d_j > 0, (j=\overline{1,n})$ and $b > 0$ are given integers.

The problem (1.1)–(1.4) is called the integer knapsack problem or the one-constrained integer programming problem in the literature [1–3], etc.

Note that, Since the problem (1)–(4) belongs to the NP-complete class, that is, to the class of "hard-to-solve problems", the maximum number of operations required by known methods (branches and bounds, dynamic programming and some combinatorial type) to find its optimal solution is exponential is from the compilation. Therefore, it is not possible to solve large-scale problems with these methods in real time. Therefore, certain approximate solution methods of problem (1.1)–(1.4) have been developed [2, 3, 6, 7], etc. These methods mainly based on the criterion

$$\max_j \frac{c_j}{a_j} = \frac{c_{j_*}}{a_{j_*}}. \tag{5}$$

So, for the number $j_*$ found from relation (5), the coordinate $x_{j_*}$ is given the maximum value that satisfies the conditions (2)–(4). Then, the next new number $j_*$ is found from relation (5) and the corresponding coordinate $x_{j_*}$ is given a value.

The process of constructing such a solution ends after evaluating all *n* number of coordinates. In this case, the value is found by looking at each coordinate only once, and the number of operations required is at most $O(n^2)$ compilation.

Without violating generality, let us assume that the coefficients of problem (1)–(4) satisfy the following relations:

$$\frac{c_1}{a_1} \ge \frac{c_2}{a_2} \ge \ldots \ge \frac{c_k}{a_k} \ge \ldots \ge \frac{c_n}{a_n}. \tag{6}$$

Then, if we ignore the completeness condition on the variables $x_j$, $(j=\overline{1,n})$ in the considered problem, the received continuous problem (1)–(3) turns into a simple linear programming problem. The optimal solution of this obtained problem is easily found analytically, and only one coordinate may not be an integer in this solution.

Suppose, this is *k*-th coordinate and is like $x_k = \frac{\alpha}{\beta}$.

We assume that the coordinates of the optimal solution of the integer knapsack problem (1)–(4) and the coordinates of the initial approximate solution differ around a certain k-th coordinate. Note that, we came to this conclusion after numerous numerical experiments, and the same result was also given in [8, 13]. If we can find those different coordinates and give a new value, it is natural that a better solution can be obtained. Such a solution algorithm is proposed in this work. For this purpose, a suitable theorem for finding the minimum number of coordinates with "0" values and values different from "0" in the optimal solution of problem (1)–(4) has been proved. In this time, we assume that the initial minimal number of non-zero coordinates and the minimal number of zeros in the easily found optimal solution of the continuous problem (1)–(3) also coincide with the optimal solution of the integer knapsack problem (1)–(4). Therefore, the initial approximate solution may have coordinates that differ from the optimal solution in that range. Because, in this range, $c_j / a_j$ ratios are not significantly different from each other. Therefore, we can get a better solution by changing the coordinates in this range within the conditions (2)–(4).

Note that such an idea was used in works [4, 5].

## 1 PROBLEM STATEMENT

**S**uppose we consider problem (1)–(4) and here relations (6) are satisfied. Then the approximate solution of that problem $X^t = (x_1^t, x_2^t, \ldots, x_n^t)$ is found by the following formula:

$$x_j^t = \begin{cases} d_j, \text{if} \quad a_j d_j \le b - \sum_{i=1}^{j-1} a_i x_i^t, \\ \left[ \left( b - \sum_{i=1}^{j-1} a_i x_i^t \right) / a_j \right], \text{if} \quad a_j d_j > b - \sum_{i=1}^{j-1} a_i x_i^t. \end{cases} \tag{7}$$


DOI 10.15588/1607-3274-2024-4-6

Here, $j=1,2,\ldots,n$ takes values and [z] is the integer part of the number $z$. Substituting this solution in the function (1), we get the approximate

$$f^t = \sum_{j=1}^{n} c_j x_j^t$$

value of that problem.

If the optimal solution of the problem (1)–(4) is $X^* = (x_1^*, x_2^*, ..., x_n^*)$, then the maximum $f^*$ value of the function (1) in this problem is as follows

$$f^* = \sum_{j=1}^{n} c_j x_j^*.$$

Obviously it must be $f^t \le f^*$. However, since finding the optimal solution $X^* = (x_1^*, x_2^*, ..., x_n^*)$ is related to serious difficulties, it is necessary to find the upper limit of the approximate value of $f^*$. Because the approximation of the approximate solution to the optimal solution should be evaluated. For this purpose, we need to solve a simple linear programming problem, ignoring the completeness condition (4) in problem (1)–(4). We call that issue an unbreakable issue. Because here the condition of being integers is not imposed on the unknowns. Note that the optimal solution of obtained unbreakable issue (1)–(3), i.e. simple linear programming problem $X^{lp} = (x_1^{lp}, x_2^{lp}, ..., x_n^{lp})$ is found by the following well-known formula.

$$x_i^{lp} = \begin{cases} d_j, \text{if } a_j d_j \le b - \sum_{i=1}^{j-1} a_i x_i^{lp}, \\ (b - \sum_{i=1}^{j-1} a_i x_i^{lp}) / a_j, \text{if } a_j d_j > b - \sum_{i=1}^{j-1} a_i x_i^{lp}, (k := j), \\ 0, \text{ if } j = k+1, k+2, ..., n. \end{cases} \quad (8)$$

So, in this solution, only the $k$-th $x_k^{lp}$ variable can take a fractional value. It is known that, if the coordinate $x_k$ is an integer, then this solution is the optimal solution of problem (1)–(4). By substituting the solution of (8) into the function (1) we will get the number

$$f^{lp} = \sum_{j=1}^{n} c_j x_j^{lp}.$$

İt is clear that, the number $f^{lp}$ is the upper limit of the approximate number $f^t$. Because we do not take into account the condition of (4) being integers over the unknowns, the set of possible solutions of the problem increases, and the maximum value obtained at this time will also be a large value. So, the following relation is true:

$$f^t \le f^* \le [f^{lp}].$$

Here, the symbol $[f^{lp}]$ indicates the integer part of the number $f^{lp}$. Thus, if we denote $[f^{lp}] = \overline{f}$, this number will be the upper limit of the approximate value of $f^t$. As a result we get $f^t \le f^* \le \overline{f}$. This relationship allow us that, to estimate the approximation of the found $f^t$ approximation to the optimal solution, that is, the absolute or relative error.

Numerous experiments and the solution of real practical problems show that the approximate solution of (7) found by the known classical way and the corresponding approximate value of $f^t$ can differ significantly from the optimal solution $X^*$ and the corresponding $f^*$ number. Therefore, our goal in performing this presented work is to develop an algorithm to find a better solution than the solution $X^t = (x_1^t, x_2^t, ..., x_n^t)$. However, this algorithm should be simple, give a greater value to the function (1) than the initial solution, should not be difficult from the programming point of view, and should be suitable for solving real practical problems. Therefore, we will call such a solution found an innovative improved approximate (suboptimal) solution.

## 2 REVIEW OF THE LITERATURE

First of all, let us note that the integer knapsack problem, including the Boolean programming problems, have been known since the last century. Since these issues have wide practical applications, various solution methods have been developed to find their optimal solutions [1–3]. Since these issues have wide practical applications, various solution methods have been developed to find their optimal solutions [1–3]. But it soon became clear that none of those methods had polynomial time complexity. In other words, the Boolean programming problem, as well as the integer bag problem, belong to the NP-complete class, that is, to the class of hard-to-solve problems [4]. Therefore, various approximate (suboptimal) solving methods were developed for this class of problems [2, 3, 5–6, etc]. On the other hand, taking into account that these issues are of wide practical importance, their more generalized models began to be applied [7–10, 14, etc.]. Here, generalization means that the given coefficients are located in certain intervals. In some works, the methods of finding the stability interval of the optimal solution and finding the generalized solution in certain problems whose coefficients are intervals have been developed [11–12 etc.].


DOI 10.15588/1607-3274-2024-4-6

Despite all this, certain studies have been conducted to further improve the approximate solutions found in integer programming problems [8, 9, 13, etc.]. In the article we have presented, the problem of finding an initial approximate solution to the integer knapsack problem has been considered.

In [15, 16, 20–23], certain approximate solution methods of knapsack problems described by various models were developed. Certain methods of solving the knapsack problem, whose initial data are in the form of intervals, are given in works [17–19]. A certain relationship between the optimal solution of the linear programming problem and the approximate solution of the integer programming problem was considered in [24], and the average case analysis of solutions of the knapsack problem with greedy algorithms was considered in [25].

## 3 MATERIALS AND METHODS

First, let's note that we can write formulas (7) and (8) more concisely as follows. Because such writing is more convenient from the point of view of programming. For each $j,(j=1,2,...,n)$ numbers

$$x_j^t = \min\left\{d_j, \left[(b - \sum_{i=1}^{j-1} a_i x_i^t)/a_j\right]\right\}, \quad (9)$$

$$x_j^{lp} = \min\left\{d_j, (b - \sum_{i=1}^{j-1} a_i x_i^t)/a_j\right\}. \quad (10)$$

Note that when finding the solution of the continuous problem (1)–(3) $X^{lp} = (x_1^{lp}, x_2^{lp}, ..., x_n^{lp})$ with the formula (10) for a certain first number $j=k$

$$(b - \sum_{i=1}^{j-1} a_i d_i)/a_j \le d_j,$$

then we remember that number $k$ and it is clear that $x_j^{lp} = 0$ for the numbers $j = k+1, k+2, ..., n.$

Thus, the optimal solution of the continuous problem (1)–(3) with the formula (10) is in the following form:

$$X^{lp} = (d_1, d_2, ..., d_{k-1}, \frac{\alpha}{\beta}, 0, ..., 0). \quad (11)$$

If the number $x_k = \frac{\alpha}{\beta}$ in the solution (11) is an integer, then this solution is the optimal solution of the integer knapsack problem (1)–(4) and no further research is needed.

Assume $x_k = \frac{\alpha}{\beta}$ is a fractional number.

Conducted numerous experiments show that the coordinates of the optimal solution $X^* = (x_1^*, x_2^*, ..., x_n^*)$ of the problem (1)–(4) are approximately $X^t = (x_1^t, x_2^t, ..., x_n^t)$ solution coordinates differ only around the $k$-th $x_k^{lp}$ coordinate in formula (11). Because, around certain $j \in [k-p; k+q]$ of the $x_k^{lp}$ coordinate, $\frac{c_j}{a_j}$ ratios do not differ significantly from each other. The selection of numbers $p$ and $q$ will be reported below.

Note that according to formula (11) $x_j^{lp} = 0$ for numbers $j = k+1, k+2, ..., k+q$ and $x_j^{lp} \ne 0$ for numbers $j = k-p, k-p+1, ..., k$. Therefore, in order to get a better solution, we can construct the new solution $X^t$ with the formula (7) by successively writing $x_j = d_j - 1, d_j - 2, ..., 0$ for each number $j = k+1, k+2, ..., k+q$ to the right of number $k$ and for number $j = k-p, k-p+1, ..., k$ to the left of number $k$. At this time, for each approximate solution of $X^t$ we calculate the new $f^t$ value of the function (1) and remember the largest value and the corresponding $X^t$ solution. We called this last-mentioned solution an innovative improved solution. Obviously, this solution will not be worse than the original solution found by the known formula (7).

It should be noted that the number p used in the interval $[k-p; k+q]$ in this article is the minimal $n(1)$ number of non-zero coordinates in the optimal solution of the problem (1)–(4) and the number q is found through the minimum $n(0)$ number of zeros in that solution. More precisely, it is chosen as $p = k - n(1), q = n - n(0) - k$. We will give the procedure for finding or evaluating the numbers and $n(0)$ below.

It should be remembered that the issue of finding a better solution by choosing a certain neighborhood of the $k$-th coordinate in the formula (11) and finding a better solution was discussed in [8,9,13]. In [13], the number $p$ was chosen as follows to determine the neighborhood of the $k$-th coordinate $[k-p, k+p]$, which received a fractional value.

$$p = \arg\left\{\max_i \frac{c_k - i}{a_k - i} - \frac{c_k + i}{a_k + i} \le \delta\right\}.$$

Here δ is the positive integer specified previously. As it can be seen, if the number $k$ is close to the last coordinates or the first coordinate of the solution (11), then it may not be possible to select the symmetric interval $[k-p, k+p]$. In the works of [8,9], only the knapsack problem with Bul variable was considered and


DOI 10.15588/1607-3274-2024-4-6

the neighborhood of the $k$-th coordinate $[\delta_k^1, \delta_k^0]$ was selected there.

Thus, $\delta_k^1 = [k \cdot \frac{q}{100}]$, $\delta_k^0 = [(n-k) \cdot \frac{q}{100}]$ are defined, and the number $q$ is the minimum number of ones or zeros in the optimal solution of the continuous knapsack problem. Here, too, the selection of the number $q$ can cause some misunderstandings. So, when the number $q$ indicates the number of units, the interval $\delta_k^1, \delta_k^0$ differs from the interval obtained when the number $q$ indicates the number of zeros.

In this work, the neighborhood of the coordinate $x_k = \frac{\alpha}{\beta}$ in the solution of (11) is chosen as $[k-p; k+q] = [n(1), n-n(0)]$.

As you can see, the relation $n(1) \le k \le n - n(0)$ is fulfilled. On the other hand, in order not to look at all the numbers located in this interval, in other words, to get the same result with a small number of operations, we need to find the minimal number $n(1)$ of non-zero coordinates and the minimal number $n(0)$ of zeros in the optimal solution of problem (1)–(4)..In this case, in the process of constructing a better solution, we should not change the first $n(1)$ number of non-zero coordinates and the last $n(0)$ number of zeros in the solution (11).

For this purpose, let's look at the following issues:

$$\sum_{j=1}^{n} x_j \to \min, \tag{12}$$

$$\sum_{j=1}^{n} a_j x_j \le b, \tag{13}$$

$$\sum_{j=1}^{n} c_j x_j \ge f^t, \tag{14}$$

$$0 \le x_j \le d_j, \ (j = \overline{1,n}), \tag{15}$$

$$x_j - \ integer \ \ (j = \overline{1,n}) \tag{16}$$

and

$$\sum_{j=1}^{n} x_j \to \max, \tag{17}$$

$$\sum_{j=1}^{n} a_j x_j \le b, \tag{18}$$

$$\sum_{j=1}^{n} c_j x_j \ge f^t, \tag{19}$$

$$0 \le x_j \le d_j, \quad (j = \overline{1,n}) \ , \tag{20}$$

$$x_j - \ integer \quad (j = \overline{1,n}). \tag{21}$$

Note that by solving problem (12)–(16), we can find the minimum n(1) number of non-zero coordinates in the optimal solution of problem (1)–(4) as number $n(\omega_1)$ is the number of elements of the set $\omega_1$.

Thus, $\omega_1 = \{j \mid \underline{x}_j^* > 0\}$, and $\underline{X}^* = (\underline{x}_1^*, \underline{x}_2^*, ..., \underline{x}_n^*)$ is the optimal solution of problem (12)–(16).

It is important to note that, the problem (12)–(16) is the special constrained integer programming problem. Obviously that, this problem from NP- complete class too and it is not easy to find their optimal solution $\underline{X}^*$. Therefore, if we do not take into account the completeness condition (16) imposed on the unknowns in that problem, the range of possible solutions of this problem will expand. Therefore, by solving the obtained linear programming problem (12)–(15), can be found the solution $\underline{X}^{lp} = (\underline{x}_1^{lp}, \underline{x}_2^{lp}, ..., \underline{x}_n^{lp})$. Then the number $n$ (1) of non-zero coordinates in this solution is found as follows:

$$\underline{n}(1) = n(\omega_2).$$

Here is $\omega_2 = \{j \mid \underline{x}_j^{lp} > 0\}$.

It is clear that, it should be $\underline{n}(1) \le n(1)$. Because, when the area grows, the minimum price can decrease. However, since finding the number $\underline{n}(1)$ is related to the solution of the linear programming problem (12)–(15) with mixed constraints, we can still encounter certain difficulties. Therefore, instead of problem (12)–(16), can be considered the following problem with a larger domain.

$$\sum_{j=1}^{n} x_j \to \min,$$

$$\sum_{j=1}^{n} c_j x_j \ge f^t,$$

$$0 \le x_j \le d_j, \ (j = \overline{1,n}) \ .$$

In this problem, without breaking generality, let us assume that the relation $c_1 \ge c_2 \ge ... \ge c_{n-1} \ge c_n$ is satisfied. Then the minimal number of non-zero coordinates $n_1$ in the optimal solution of this problem can be found as follows:

$$\sum_{j=1}^{n_1} c_j d_j \le f^t \le \sum_{j=1}^{n_1+1} c_j d_j \ .$$

It is clear that, the relationship $n_1 \le \underline{n}(1) \le n(1)$ is satisfied.


DOI 10.15588/1607-3274-2024-4-6

Thus, we can take the number $n_1$ which is easily found, as the number of coordinates different from zero in the optimal solution of the problem (1)–(4).

Now let's find the minimum number of zeros in the optimal solution of problem (1)–(4). For this purpose, by making judgments according to the above, we have to solve the problem (17)–(21) or the problem (17)–(20), and finally the problem (22)–(25):

$$\sum_{j=1}^{n} x_j \to \max, \tag{22}$$

$$\sum_{j=1}^{n} a_j x_j \le b, \tag{23}$$

$$0 \le x_j \le d_j,\ (j=\overline{1,n}), \tag{24}$$

$$x_j - \ integer \quad (j=\overline{1,n}). \tag{25}$$

In the case of (22)–(25), let us assume, without violating the generality, that the relations $a_1 \le a_2 \le ... \le a_n$ are fulfilled in the condition (23). Then the maximal number $\bar{n}$ of non-zero coordinates in the optimal solution of this problem is found from the following relationship:

$$\sum_{j=1}^{\bar{n}} a_j d_j \le b \le \sum_{j=1}^{\bar{n}+1} a_j d_j.$$

It is clear that, we can take the number $\underline{n}(0) = n - \bar{n}$ instead of the minimal $n(0)$ number of zeros in the optimal solution of problem (1)–(4).

Thus, we proved the following theorem.

**Theorem:** The inequalities $\underline{n}(1) \le n(1)$ and $\underline{n}(0) \le n(0)$ for the minimum $n(1)$ number of non-zero coordinates and the minimum $n(0)$ number of zeros in the optimal solution of problem (1)–(4) it is true.

Let's note that the coordinates of the optimal solution of the problem (1)–(4) with the solution (11) can differ in the interval $[\underline{n}(1), n - \underline{n}(0)]$.

It is clear that $d_j$ numbers are to the left and only zeros are to the right of the k-th coordinate in the solution of (3.3). Therefore, for each number $j$, ($j = \underline{n}(1), \underline{n}(1)+1, ..., k$) to the left of the $x_j = 1,2,...,d_j$, $x_j = d_{j-1}, d_{j-2}, ..., 0$ and $x_j = 1,2,...,d_j$ for each ($j$=$k$+1, $k$+2,...) numbers on the right, new solutions can be constructed by formulas (7) or (9). We will select the best of those solutions and consider it as an innovative improved approximate solution.

Now, let's write the algorithm for the innovative approximate solution method described above.

## ALGORITHM

**Step 1.** Enter the numbers $n, b, c_j, a_j, d_j\ (j=\overline{1,n})$ and accept $bb := b$;

**Step 2.** For each number $j, (j = 1,2,...,n)$ you need to find the solution $X^{lp} = (x_1^{lp}, x_2^{lp}, ..., x_k^{lp}, ..., x_n^{lp})$ with the formula

$$x_j^{lp} = \begin{cases} d_j, \ if \ \ a_j d_j \le b - \sum_{i=1}^{j-1} a_i x_i^{lp}, \\ (b - \sum_{i=1}^{j-1} a_i x_i^{lp}) / a_j, \quad if \quad a_j d_j > b - \sum_{i=1}^{j-1} a_i x_i^{lp}, \ (k := j), \\ 0, \quad when \quad j = k+1,\ k+2, ..., n \end{cases}$$

and from here the fractional coordinate number $k$ should be noted. kk:=k; r:=0;

**Step 3.** If $x_k^{lp}$ is an integer, then the solution $X^{lp}$ is the optimal solution of problem (1)–(4). At this time we should calculate

$$f^* := \sum_{j=1}^{n} c_j x_j^{lp},$$

$X^* = (x_1^*, x_2^*, ..., x_n^*) = (x_1^{lp}, x_2^{lp}, ..., x_n^{lp})$, print $f^*$ and go to step 21.

**Step 4.** To find the approximate solution of $X^t = (x_1^t, x_2^t, ..., x_n^t)$ we need to calculate

$$x_j^t = \begin{cases} d_j, \ if \ \ a_j d_j \le b - \sum_{i=1}^{j-1} a_i x_i^t, \\ [(b - \sum_{i=1}^{j-1} a_i x_i^t) / a_j], \quad if \quad a_j d_j > b - \sum_{i=1}^{j-1} a_i x_i^t, \end{cases}$$

for each number $(j = 1,2,...,n)$.

**Step 5.** Calculate the numbers

$$f^{lp} = \sum_{j=1}^{n} c_j x_j^{lp},\ f^t = \sum_{j=1}^{n} c_j x_j^t.$$

Accept $f^{it} := f^t$ and remembered the solution $X^{it} = (x_1^t, x_2^t, ..., x_n^t)$ with $f^{it}$.

**Step 6**. Coefficients $c_j, \left(j=\overline{1,n}\right)$ should be arranged as $c_1 \ge c_2 \ge ... \ge c_n$ and to find the minimal $n_1$ number of non-zero coordinates in the optimal solution from the relationship


DOI 10.15588/1607-3274-2024-4-6

$$\sum_{j=1}^{n_1} c_j d_j \le f^t \le \sum_{j=1}^{n_1+1} c_j d_j .$$

**Step 7**. Numbers $a_j, (j=\overline{1,n})$ should be arranged as $a_1 \le a_j \le ... \le a_n$ and find the maximum $\bar{n}$ number of coordinates different from zero in the optimal solution of the problem (1)–(4) from the relationship and note $n(0) = n - \bar{n}$

$$\sum_{j=1}^{\bar{n}_1} a_j d_j \le b \le \sum_{j=1}^{\bar{n}_1+1} a_j d_j$$

and note $\underline{n}(0) = n - \bar{n}$.

**Step 8.** Set $x_k^t := [x_k^t];\ b := bb - a_k \times x_k^t$.

**Step 9**. For the numbers $j,\ j=1,2,...,n\quad j \ne k$

$$x_j^t = \begin{cases} d_j,\ if\ a_j d_j \le b - \sum_{i=1}^{j-1} a_i x_i^t, \\ [(b - \sum_{i=1}^{j-1} a_i x_i^t)/a_j],\quad if\quad a_j d_j > b - \sum_{i=1}^{j-1} a_i x_i^t. \end{cases}$$

**Step 10**. Calculate

$$f^t := \sum_{j=1}^{n} c_j x_j^t .$$

If $f^t > f^{it}$, then should be accept $f^{it} := f^t$, $X^{it} = (x_1^t, x_2^t, ..., x_n^t)$.

**Step 11**. If $r$=0, then should be accept $x_k^t := [x_k^t] + 1;\ b := bb - a_k \times x_k^t;\ r := 1$ and go to step 9.

**Step 12** Set be accept $r := 0$

**Step 13**. For each $j$, ( $j = 1,2,...,n;\ j \ne k$ )

$$x_j^t = \begin{cases} d_j,\ if\ a_j d_j \le b - \sum_{i=1}^{j-1} a_i x_i^t, \\ [(b - \sum_{i=1}^{j-1} a_i x_i^t)/a_j],\quad if\quad a_j d_j > b - \sum_{i=1}^{j-1} a_i x_i^t. \end{cases}$$

With the formula $X^t = (x_1^t, x_2^t, ..., x_n^t)$ calculate the number $f^t$

$$f^t := \sum_{j=1}^{n} c_j x_j^t .$$

**Step 14.** If $f^t > f^{it}$, ther $f^{it} := f^t$, $X^{it} = (x_1^t, x_2^t, ..., x_n^t)$ should be written and memorized. If $r := 1$ go to Step 18.

**Step 15**. If $x_k^t < d_k$, then go to Step 17.

**Step 16**. $k := k+1$; If $k > n(0)$ $k := kk$ and go to the Step 18.

**Step 17**. $x_k^t := 1,2,...,d_k$ and accordingly by taking $b := bb - a_k \times x_k^t$ go to the Step 13.

**Step 18**. $r := 1$; If $x_k^t = d_k$ , then $k := k-1$; If $k < n(1)$ go to the Step 20.

**Step 19**. $x_k^t := 1,2,...,d_k$ $k$ values corresponding to $b := bb - a_k \times x_k^t$ and go to Step 13.

**Step 20**. Print $f^{it}$, $X^{it} = (x_1^{it}, x_2^{it}, ..., x_n^{it})$. $\delta = ([f^{lp}] - f^t / [f^{lp}]$ and $\delta^i = ([f^{lp}] - f^{it} / [f^{lp}]$.

**Step 21**. STOP.

## 4 EXPERIMENTS

Numerous computational experiments have been conducted to investigate the quality of the innovative improved solution method we proposed above.

During the experiments, problems with a different number of variables were solved ($n$=100, $n$=300, $n$=500, $n$=1000). The coefficients of these problems were chosen as random numbers with two digits and three digits at most.

Thus, $0 < c_j \le 99,\ 0 < a_j \le 99\ d_j = 10, (j = \overline{1,n})$ or $0 < c_j \le 999,\ 0 < a_j \le 999\ d_j = 10, (j = \overline{1,n})$,

$$b = [\frac{1}{3} \sum_{j=1}^{n} a_j d_j].$$

It should be noted that 4 different problems, each with the same number of variables, were solved.

The results are given in the following tables and notations are adopted as follows.

$N$ – the number of the solved problem with the same number of unknowns.

$\bar{f}$ – the upper bound of the optimal value of the problem (1.1)–(1.4).

$f^t$ – the value of the function (1.1) according to the approximate solution of the problem (1.1)–(1.4) found by the known classical method.

$f^{it}$ – the value given to the function (1.1) of the improved approximate solution.

$\delta$ – the relative error of the approximate value found by the classical method, expressed as a percentage of the optimal value. It mean that,

$$\delta = \frac{\bar{f} - f^t}{\bar{f}} \cdot 100 ,$$


DOI 10.15588/1607-3274-2024-4-6

$\delta^i$ – an innovative improved approximate persentage is relative error.

$$\delta^i = \frac{\overline{f} - f^i}{\overline{f}} \cdot 100 .$$

## 5 RESULTS

Table 1 – Problems with two-digit coefficients ($n = 100$)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 34211.00 | 34493.00 | 37378.00 | 33508.00 |
| $f^t$ | 34205.00 | 34486.00 | 37353.00 | 33492.00 |
| $f^{it}$ | 34209.00 | 34491.00 | 37376.00 | 33506.00 |
| $\delta$ | 0.01754 | 0.02029 | 0.06688 | 0.04775 |
| $\delta^i$ | 0.00585 | 0.00580 | 0.00535 | 0.00597 |

Table 2 – Problems with two-digit coefficients ($n = 300$)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 103013.00 | 98887.00 | 97245.00 | 99692.00 |
| $f^t$ | 103010.00 | 98884.00 | 97244.00 | 99684.00 |
| $f^{it}$ | 103012.00 | 98885.00 | 97246.00 | 99693.00 |
| $\delta$ | 0.00291 | 0.00303 | 0.00103 | 0.00802 |
| $\delta^i$ | 0.00097 | 0.00202 | 0.00103 | 0.00100 |

Table 3 – Problems with two-digit coefficients ($n$=500)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 161131.00 | 171036.00 | 170555.00 | 168122.00 |
| $f^t$ | 161128.00 | 171033.00 | 170555.00 | 168115.00 |
| $f^{it}$ | 161130.00 | 171034.00 | 170555.00 | 168122.00 |
| $\delta$ | 0.00186 | 0.00175 | 0.00000 | 0.00416 |
| $\delta^i$ | 0.00062 | 0.00117 | 0.00000 | 0.00000 |

Table 4 – Problems with two-digit coefficients ($n$=1000)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 339009.00 | 330007.00 | 329556.00 | 335835.00 |
| $f^t$ | 339009.00 | 330003.00 | 329555.00 | 335832.00 |
| $f^{it}$ | 339009.00 | 330007.00 | 329556.00 | 335834.00 |
| $\delta$ | 0.00000 | 0.00121 | 0.00030 | 0.00089 |
| $\delta^i$ | 0.00000 | 0.00000 | 0.00000 | 0.00030 |

Table 5 – Problems with three-digit coefficients ($n = 100$)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 343742. | 346062.00 | 375258.0 | 336331.00 |
| $f^t$ | 343485.00 | 345998.00 | 375018.00 | 336155.00 |
| $f^{it}$ | 343660.00 | 346040.00 | 375227.00 | 336305.00 |
| $\delta$ | 0.07477 | 0.01849 | 0.06396 | 0.05233 |
| $\delta^i$ | 0.02386 | 0.00636 | 0.00826 | 0.00773 |

Table 6 – Problems with three-digit coefficients ($n = 300$)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 1035261.00 | 993336.00 | 976972.00 | 1002027.00 |
| $f^t$ | 1035236.00 | 993265.00 | 976948.00 | 1001939.00 |
| $f^{it}$ | 1035251.00 | 993307.00 | 976948.00 | 10019 95.00 |
| $\delta$ | 0.00241 | 0.00715 | 0.00246 | 0.00878 |
| $\delta^i$ | 0.00097 | 0.00292 | 0.00246 | 0.00319 |

Table 7 – Problems with three-digit coefficients ($n$=500)

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\overline{f}$ | 1620326.00 | 1718600.00 | 1712542.00 | 1689749.00 |
| $f^t$ | 1620292.00 | 1718561.00 | 1712502.00 | 1689670.00 |
| $f^{it}$ | 1620303.00 | 1718587.00 | 1712528.00 | 1689739.00 |
| $\delta$ | 0.00210 | 0.00227 | 0.00234 | 0.00468 |
| $\delta^i$ | 0.00142 | 0.00076 | 0.00082 | 0.00059 |

## 6 DISCUSSION

Based on the tables, the following conclusions can be drawn.

In most cases, the initially found approximate solution has been further improved. Rather, in 24 out of 28 solved problems, the initial approximate solution was further improved. In the remaining 4 problems, the initial approximate solution has not improved. It can be assumed that this solution is the optimal solution. The relative errors of the found approximate values from the optimal value are very small and do not exceed 1%. This is very important for solving real practical problems. It should be noted that the algorithm proposed in the article does not count options, so it takes seconds to solve problems. Therefore, we did not mention the computer time in the tables.

## CONCLUSIONS

A new approximate solution method of the integer knapsack problem is given in the presented article. Through this method, any initial solution found by known methods is successively improved. At this time, the


DOI 10.15588/1607-3274-2024-4-6

interval where the coordinates that do not coincide with the optimal solution are located is determined. After that, it is possible to build a better solution by assigning new values to the coordinates in those intervals.

The proposed method is new, simple in nature, easy to consider from the programming point of view, and has important practical importance, so this method is called an innovative improved approximate solution method.

## ACKNOWLEDGEMENTS

This work, i.e. “Innovative improved approximate solution method for the integer kapsack problem, error compression and computational experiments” was carried out at the Institute of Control Systems of the Ministry of Science and Education of the Republic of Azerbaijan at the expense of the state budget (State Registration No. 0101 Az 00736). Note that brief information about this work was given for the first time in [8].

## REFERENCES

1. Erlebach T. Kellerer H., Pferschy U. Approximating multi-objective knapsack problems, *Management Science*, 2002, № 48, pp. 1603–1612. DOI: 10.1287/mnsc.48.12.1603.445
2. Kellerer H., Pferschy U., Pisinger D. Knapsack problems. Berlin, Heidelberg, New-york, Springer- Verlag, 2004, P. 546. DOI:10.1007/978-3-540-24777-7
3. Martello S., Toth P. Knapsack problems: Algorithm and Computers İmplementations. New York, John Wiley & Sons, 1990, P. 296 .
4. Garey M. R., Jhonson D. S. Computers and Intractability : a Guide to the Theory of NP-Completeness. San Francisco, Freeman, 1979, P. 314.
5. Vazirani V. V. Approximation algorithms. Berlin, Springer, 2001, p. 378. DOI:10.1057/palgrave.jors.2601377
6. Bukhtoyarov S. E., Emelichev V. A. Stability aspects of Multicriteria integer linear programming problem, *Journal of Applied and Industrial mathematics*, 2019, V(13), №1, pp. 1–10. DOI:10.33048/daio.2019.26.624
7. Libura M. Integer programming problems with inexact objective function, *Control Cybern*, 1980, Vol. 9, N 4, pp. 189–202.
8. Niyazova R. R., Huseynov S. Y. An İnnovative İmproved Approximate Method for the knapsack Problem with coefficients Given in the Interval Form, *8-th International Conference on Control and Optimization with Industrial Applications*. Baku, 24 – 26 August, 2022. Vol II, pp. 210 – 212.
9. Mammadov K. Sh., Niyazova R. R.,. Huseynov S. Y. Innovative approximate method for solving Knapsack problems with interval coefficients, *International Independent scientific journal*, 2022, № 44, pp. 8–12. doi.org/10.5281/zenodo.7311206.
10. Mamedov K. SH., Mammadli N. O. Two methods for construction of suboptimistic and subpessimistic solutions of the interval problem of mixed-Boolean programming, *Radio Electronics, Computer Science, Control*, 2018, № 3 (46), pp. 57–67. DOI 10.15588/1607-3274-2018-3-7.
11. Emelichev Vladimir, Podkopaev Dmitry Quantitative stability analysis for vector problems of 0–1 programming, *Discrete Optimization*, 2010, vol. 7, pp. 48–63. DOI: 10.1016/j.disopt.2010.02.001.
12. Li W. Liu X., Li H. Generalized solutions to interval linear programmers and related necessary and sufficient optimality conditions, *Optimization Methods Software*, 2015, Vol. 30, №3, pp. 516–530 .DOI: 10.1080/10556788.2014.940948
13. Mamedov K. Sh., Huseinov S. Y. Method of Constructing Suboptimal Solutions of Integer Programming Problems and Successive Improvement of these Solutions, *Automatic Control and Computer Science*, 2007, Vol. 41, № 6, pp. 312–319. DOI: 10.3103/S014641160706003X
14. Mamedov K. Sh., Mamedova A. H. Ponyatie suboptimisticheskoqo i subpestimisticheskoqo resheniy i postroeniya ix v intervalnoy zadache Bulevoqo programmirovania, Radioelektronika, Informatika, Upravlenie, 2016, N3, pp. 99–108. DOİ: 10.15588/1607-3274-2016-3-13
15. Hifi M., Sadfi S., Sbihi A. An efficient algorithm for the knapsack sharing problem, *Computational Optimization and Applications*, 2002, № 23, pp. 27–45. DOI:10.1023/A:1019920507008
16. Hifi M., Sadfi S. The knapsack sharing problem: An exact algorithm, *Journal of Combinatorial Optimization*, 2002, № 6, p. 35–54. DOI:10.1023/A:1013385216761
17. Hladik M. On strong optimality of interval linear programming, *Optimization Letters*, 2017, Vol. 11(7), pp. 1459–1468. DOI:10.1007/s11590-016-1088-3
18. Devyaterikova M. V., Kolokolov A. A. L-class enumeration algorithms for knapsack problem with interval data, *International Conference on Operations Research: Book of Abstracts*. Duisburg, 2001, P. 118.
19. Devyaterikova M. V., Kolokolov A. A., Kolosov A. P. L-class enumeration algorithms for one discrete production planning problem with interval input data, *Computers and Operations Research*, 2009, Vol. 36, №2, pp. 316–324. DOI:10.1016/j.cor.2007.10.005
20. Babayev D. A., Mardanov S. S. Reducing the number of variables in integer and linear programming problems, *Computational Optimization and Applications*, 1994, № 3, pp. 99–109. DOI: https://doi.org/10.1007/BF01300969.
21. Basso A., Viscolani B. Linear programming selection of internal financial laws and a knapsack problem, *Calcolo*, 2000, V.37, № 1, pp. 47–57. DOI:10.1007/s100920050003
22. Bertisimas D., Demir R. An approximate dynamic programming approach to multidimensional knapsack problems, *Management Science*, 2002, № 48, pp. 550–565. DOI: 10.1287/mnsc.48.4.550.208
23. Billionnet A. Approximation algorithms for fractional knapsack problems, *Operation Research Letters*, 2002, № 30, pp. 336–342. DOI:10.1016/S0167-6377(02)00157-8
24. Broughan Kevin, Zhu Nan An integer programming problem with a linear programming solution, *Journal American Mathematical Monthly*, 2000, V.107, № 5, pp. 444–446. DOI: 10.1080/00029890.2000.12005218
25. Calvin J. M., Leung J. Y-T. Average-case analysis of a greedy algorithm for the 0–1 knapsack problem, *Operation Research Letters*, 2003, № 31, pp. 202–210. DOI: 10.1016/S0167-6377(02)00222-5

Received 24.07.2024.
Accepted 27.10.2024.


DOI 10.15588/1607-3274-2024-4-6

УДК 519.852.6

# ІННОВАЦІЙНИЙ ВДОСКОНАЛЕНИЙ МЕТОД НАБЛИЖЕНОГО РІШЕННЯ ДЛЯ ЗАДАЧІ ЦІЛОЧИСЕЛЬНОГО РАНЦЯ, СТИСНЕННЯ ПОМИЛОК ТА ОБЧИСЛЮВАЛЬНІ ЕКСПЕРИМЕНТИ

**Мамедов К. Ш. –** д-р фіз.-мат. наук, професор Бакинського державного університету та завідувач відділу Інституту систем управління Міністерства науки і освіти.

**Ніязова Р. Р.** – докторант, науковий співробітник Інституту систем управління Міністерства освіти і науки.

## АНОТАЦІЯ

**Актуальність**. Математичні моделі багатьох задач оптимізації, що зустрічаються в економіці та техніці, розглядаються у формі задачі про цілочисельний рюкзак. Оскільки ця задача належить до класу «NP-повних», тобто «важко розв'язуваних», кількість операцій, необхідних відомим методам для знаходження її оптимального розв'язку, експоненціальна. Це не дозволяє вирішувати масштабні завдання в режимі реального часу. Тому розроблено різноманітні та швидкопрацюючі методи наближеного розв'язання цієї задачі. Однак відомо, що наближене рішення, отримане цими методами, у більшості випадків може суттєво відрізнятися від оптимального. Тому після прийняття будь-якого наближеного рішення за вихідну точку виникає потреба розробити методи його подальшого вдосконалення. Розробка таких методів має як теоретичне, так і велике практичне значення.

**Мета роботи**. Основна мета вирішення цього питання полягає в наступному. Основна мета виконання даної роботи полягає в тому, щоб будь-яким відомим методом спочатку знайти вихідний наближений розв'язок задачі, а потім розробити алгоритм для послідовного подальшого вдосконалення цього розв'язку. Для цього необхідно визначити набір чисел, якими можуть відрізнятися координати оптимального і знайденого наближеного розв'язку. Після цього слід побудувати нові розв'язки шляхом присвоєння можливих значень невідомим, що відповідають числам цього набору, і вибрати найкраще з цих розв'язків. Але алгоритм побудови такого рішення повинен бути простим, вимагати невеликої кількості операцій, не викликати труднощів з точки зору програмування, бути новим і застосовним до практичних завдань.

**Метод.** Суть запропонованого способу полягає в наступному. Спочатку за відомим правилом знаходять початковий наближений розв'язок задачі, що розглядається, і відповідне йому значення цільової функції. Після цього оптимальний розв'язок задачі легко знаходить відомим методом без урахування умови цілості невідомих. Очевидно, що цей розв'язок може приймати не більше одного дробового значення координати. Передбачається, що координати оптимального розв'язку цілочисельної задачі про ранець і початкового наближеного розв'язку можуть відрізнятися навколо певної дробової координати оптимального розв'язку неперервної задачі. Потім знайдено мінімальну кількість ненульових координат і нульових координат в оптимальному розв'язку. Для цього доведено відповідні теореми. Передбачається, що різні координати оптимального розв'язку та початкового наближеного розв'язку знаходяться між цими мінімальними числами. Таким чином, найкраще рішення можна вибрати шляхом послідовної зміни координат між цими мінімальними числами один за одним.

**Результати.** Із застосуванням запропонованого методу були проведені численні розрахункові експерименти. Висока якість цього методу ще раз підтверджена експериментально.

**Висновки.** Запропонований метод є новим, простим за своєю суттю, легким для програмування та має важливе практичне значення. Таким чином, ми називаємо це рішення інноваційним покращеним наближеним рішенням.

**КЛЮЧОВІ СЛОВА:** задача про цілочисельний ранець, розв'язок початкового наближення, мінімальна кількість нульових і ненульових координат в оптимальному розв'язку, інноваційне вдосконалення розв'язку початкового наближення, оцінка похибки та експерименти.

## ЛІТЕРАТУРА

1. Erlebach T. Approximating multi-objective knapsack problems / T. Erlebach, H. Kellerer, U. Pferschy // Management Science. – 2002. – № 48. – P. 1603–1612. DOI: 10.1287/mnsc.48.12.1603.445
2. Kellerer H. Knapsack problems. / H. Kellerer, U. Pferschy, D. Pisinger. – Berlin, Heidelberg, New-york : Springer-Verlag, 2004. – P. 546. DOI:10.1007/978-3-540-24777-7
3. Martello S. Knapsack problems: Algorithm and Computers İmplementations. / S. Martello, P. Toth. – New York : John Wiley & Sons, 1990. – P. 296.
4. Garey M. R. Computers and Intractability : a Guide to the Theory of NP-Completeness. / M. R. Garey, D. S. Jhonson. – San Francisco, Freeman, 1979. – P. 314.
5. Vazirani V. V. Approximation algorithms / V. V. Vazirani. – Berlin : Springer, 2001. – P. 378. DOI:10.1057/palgrave.jors.2601377
6. Bukhtoyarov S. E. Stability aspects of Multicriteria integer linear programming problem / S. E. Bukhtoyarov, V. A. Emelichev // Journal of Applied and Industrial mathematics. – 2019. – V(13), № 1. – P. 1–10. DOI:10.33048/daio.2019.26.624
7. Libura M. Integer programming problems with inexact objective function / M. Libura // Control Cybern. – 1980. – Vol. 9, N 4. – P. 189–202.
8. Niyazova R. R. An İnnovative İmproved Approximate Method for the knapsack Problem with coefficients Given in the Interval Form / R. R. Niyazova , S. Y. Huseynov // 8-th International Conference on Control and Optimization with Industrial Applications, Baku : 24–26 August, 2022. Vol II. –P. 210–212.
9. Mammadov K. Sh. Innovative approximate method for solving Knapsack problems with interval coefficients. / K. Sh. Mammadov, R. R. Niyazova, S. Y. Huseynov // International Independent scientific journal. – 2022, № 44. – P. 8–12. doi.org/10.5281/zenodo.7311206.
10. Mamedov K. Sh. Two methods for construction of suboptimistic and subpessimistic solutions of the interval problem of mixed-Boolean programming / K. Sh. Mamedov, N. O. Mammadli // Radio Electronics, Computer Science, Control. – 2018, № 3(46). – P. 57–67. DOI 10.15588/1607-3274-2018-3-7.
11. Vladimir Emelichev Quantitative stability analysis for vector problems of 0–1 programming / Vladimir Emelichev,


DOI 10.15588/1607-3274-2024-4-6

Dmitry Podkopaev // Discrete Optimization. – 2010. – vol. 7. – P. 48–63. DOI: 10.1016/j.disopt.2010.02.001.
12. Li W. Generalized solutions to interval linear programmers and related necessary and sufficient optimality conditions / W. Li, X. Liu, H. Li // Optimization Methods Software. – 2015, Vol. 30, № 3. – P. 516–530. DOI: 10.1080/10556788.2014.940948
13. Mamedov K. Sh. Method of Constructing Suboptimal Solutions of Integer Programming Problems and Successive Improvement of these Solutions / K. Sh. Mamedov, S. Y. Huseinov // Automatic Control and Computer Science. – 2007. – Vol. 41, № 6. – P. 312–319. DOI: 10.3103/S014641160706003X
14. Mamedov K. Sh. Ponyatie suboptimisticheskoqo i subpestimisticheskoqo resheniy i postroeniya ix v intervalnoy zadache Bulevoqo programmirovania / K. Sh. Mamedov, A. H. Mamedova // Radioelektronika, Informatika, Upravlenie. – 2016. – No. 3. – P. 99–108. DOİ: 10.15588/1607-3274-2016-3-13
15. Hifi M. An efficient algorithm for the knapsack sharing problem / M. Hifi, S. Sadfi, A. Sbihi //.Computational Optimization and Applications. – 2002. – № 23. – P. 27–45. DOI:10.1023/A:1019920507008
16. Hifi M. The knapsack sharing problem: An exact algorithm. / M. Hifi, S. Sadfi // Journal of Combinatorial Optimization. – 2002. – № 6. – P. 35–54. DOI:10.1023/A:1013385216761
17. Hladik M. On strong optimality of interval linear programming / M. Hladik // Optimization Letters. – 2017. – Vol. 11(7). – P. 1459–1468. DOI:10.1007/s11590-016-1088-3
18. Devyaterikova M. V. L-class enumeration algorithms for knapsack problem with interval data / M. V. Devyaterikova, A. A. Kolokolov // International Conference on Operations Research: Book of Abstracts. – Duisburg, 2001. – P. 118.
19. Devyaterikova M. V. L-class enumeration algorithms for one discrete production planning problem with interval input data / M. V. Devyaterikova, A. A. Kolokolov, A. P. Kolosov // Computers and Operations Research. –2009. – Vol. 36, № 2. – P. 316–324. DOI:10.1016/j.cor.2007.10.005
20. Babayev D. A. Reducing the number of variables in integer and linear programming problems / D. A. Babayev, S. S. Mardanov // Computational Optimization and Applications. – 1994. – № 3. – P. 99–109. DOI: https://doi.org/10.1007/BF01300969.
21. Basso A. Linear programming selection of internal financial laws and a knapsack problem/ A. Basso, B. Viscolani // Calcolo. – 2000. – V. 37, № 1. – P. 47–57. DOI:10.1007/s100920050003
22. Bertisimas D. An approximate dynamic programming approach to multidimensional knapsack problems. / D. Bertisimas, R. Demir // Management Science. – 2002. – № 48. – P. 550–565. DOI:10.1287/mnsc.48.4.550.208
23. Billionnet A. Approximation algorithms for fractional knapsack problems / A. Billionnet // Operation Research Letters. – 2002. – № 30. – P. 336–342. DOI:10.1016/S0167-6377(02)00157-8
24. Broughan Kevin An integer programming problem with a linear programming solution / Kevin Broughan, Nan Zhu // Journal American Mathematical Monthly. – 2000. – V. 107, № 5. – P. 444–446. DOI:10.1080/00029890.2000.12005218
25. Calvin J. M. Average-case analysis of a greedy algorithm for the 0–1 knapsack problem / J. M. Calvin, J. Y-T. Leung // Operation Research Letters. – 2003. – № 31. – P. 202–210. DOI:10.1016/S0167-6377(02)00222-5


DOI 10.15588/1607-3274-2024-4-6

UDC 519.85

# SOLUTION OF A MULTICRITERIA ASSIGNMENT PROBLEM USING A CATEGORICAL EFFICIENCY CRITERION

**Novozhylova M. V.** – Dr. Sc., Professor, Head of the Department of Computer Science and Information Technologies, O. M. Beketov National University of Urban Economy in Kharkiv, Ukraine.

**Karpenko M. Yu**. – PhD, Associate Professor of the Department of Computer Science and Information Technologies, O. M. Beketov National University of Urban Economy in Kharkiv, Kharkiv, Ukraine.

## ABSTRACT

**Context.** The paper considers a problem of assigning a set of employees to a finite set of operations in a multicriteria statement, under condition of a hierarchical structure of a partial efficiency criterion of performing a set of operations, being presented in such a way that each employee possesses a finite set of competencies and each operation has a finite set of characteristics. Numerical and categorical data types are provided for the use as exogenous parameters of the problem. The relevance of the assignment problem being considered is determined by an extremely wide range of practical applications, both in the classical statements and new modifications, the high demand for which is constantly generated by the dynamically developing economic environment. At the same time, a critically smaller number of scientific publications propose means of modeling and solving multi-criteria assignment problems, despite the importance of this type of problems in decision-making, both in theoretical and practical aspects. In general, in conditions of lack of information, the exogenous parameters of the problem cannot be specified in numerical form, therefore there is a need to use categorical data with further numerical coding.

**Objective.** The goal of the work is to build a multicriteria mathematical model and, on this basis, carry out a numerical study of the optimization assignment problem, taking into account a hierarchical structure of a partial efficiency criterion of the selection of «operation – employee» pairs.

**Method.** The study proposes a novel method of solving the assignment problem that implemented as a multi-stage process, which includes the stage of transformation of exogenous parameters of the model, given by categorical variables, based on the implementation of the Pareto principle and logistic mapping, the stage of constructing linear scalarization of the efficiency and the cost criteria.

**KEYWORDS:** mathematical and computer modeling of an assignment problem, multicriteria optimization, Pareto set, categorical parameters, logistic curve.

## ABBREVIATIONS

NP-complexity is an nondeterministic polynomial complexity;

DM is a decision maker.

## NOMENCLATURE

$i$ is an index of operation;

$j$ is an index of employee;

$k$ is an index of operation characteristics;

$I$ is a number of operations;

$J$ is a number of employees;

$K$ is a number of operation characteristics;

$A$ is a matrix of weighting coefficients;

$B^i$ is a binary $K\times 3$ matrix;

$b_{kn}^i$ is an element of binary matrix $B^i$;

$c_{ij}$ is a cost of performing the $i$-th operation by the $j$-th employee;

$C$ is a cost matrix;

$C^H$ is a normalized cost matrix;

$c_{ij}^{norm}$ is an element of the normalized cost matrix;

$e_{ii}(v_i, s_j)$ is a similarity measure of vectors $(v_i, s_j)$ ;

$c^{\min}$ the minimum value of cost matrix elements;

$c^{\max}$ the maximum value of cost matrix elements;

$E$ is a performance effectivity matrix;

$G(C,X)$ is a cost criterion of the problem;

$f(M)$ is a function of digitalization of categorical variables;

$K_E$ is an effectivity function;

$\hat{K}_E$ is an approximation of an effectivity function;

$K_C$ is a cost function;

$\boldsymbol{M}$ is a set of categorical values**;**

$N_1$ is a value from the interval (0,1);

$N_2$ is a value from the interval (0,1), $N_1 < N_2$;

$N$ is a number of labels in dummy coding;

$R^2$ is a reliability of approximation;

$s_i$ is a vector of competence levels of an $j$-th employee;

$s_{jk}$ is a level of competence of an $j$-th employee regarding $k$-th characteristic of operation;

$S$ is a matrix of levels of employees' competencies;

$x_{ii}$ is a binary endogenous variable of the assignment problem;

$X$ is a matrix of endogenous variables of the assignment problem;

$v_i$ is a vector of operation $\theta_i$ characteristics;

$v_{ik}$ is a $k$-th characteristic of operation $\theta_i$;

$w_i$ is a ordered set of operation $\theta_i$ characteristics;

$\alpha_i$ is a assessment vector of operation $\theta_i$ characteristics;

$\alpha_{ik}$ is a $k$-th component of the vector $\alpha_i$;

$\Theta$ is a set of operations;

$\theta_i$ is an $i$-th operation;

$\Xi$ is a team of employees;

$\omega_j$ is a $j$-th employee;


DOI 10.15588/1607-3274-2024-4-7

$\lambda_1$, $\lambda_2$ is a weighting coefficients of partial criteria in a assignment problem;

$\Phi(E,X)$ is an efficiency criterion to select optimal pairs $(\theta_i, \omega_j)$;

$\times$ is a cartesian product of sets.

## INTRODUCTION

Assignment problems as a scientific direction propose instrumental tools to optimal use of production capacities and resources for the livelihood of consumers that is one of the urgent problems today. To solve this problem is even more important when functioning of enterprises is under pressure of hostile external environment and all the production and economic activities are forced to take place in the conditions of an emergency situation, and the conduct of continued military operations.

Thus assingment problems both in the classical statement and in new modifications belong to the most widespread optimization problems of resource distribution, which arise in business practice, in particular in the urban economy of large cities, as well as when providing services to the population, individuals and legal entities in the field of retail, medical care, hotel business, transport and repair services, supply organization, etc., and the demand for which is constantly generated by life itself.

The obvious practical and theoretical importance of the problem for the development of the theory of algorithms and the theory of optimization in general at least because a quadratic assignment problem belongs to the class of combinatorial optimization problems and it is NP-complex – causes continiuos interest of scientists.

Without losing generality let's consider this problem in the statement about the assignment of employeers having various qualifications for operational, repairing, monitoring activities that relate to one or a group of complex technical or social objects.

Thus, let there be a set $\Theta = \{\theta_i\}$, $i$=1,...,$I$, of certain operations (objects) and a team (set) $\Xi = \{\omega_j\}$, $j$=1,...,$J$, of employers (resources), in general $I \neq J$.

We need to assign at least one operation to each $j$-th employee, and only one employee to each $i$-th operation. For each such pair $(\theta_i, \omega_j)$ the cost $c_{ij}$ of performing an operation $\theta_i$ by an employer $\omega_j$ is determined. According to the classical formulation, the objective function models the total costs. It is necessary to find the minimum value of objective function.

Assignment problems admit of several generalizations. We will further consider multicriteria problems as a natural generalization of the classical statement. In multicriteria problems, a partial cost criterion is considered along with a partial criterion of the effectivity of operation performance [1].

Thus, along with the cost of performing operations, the effectivity function of performing operations is also considered, which, depending on the specific statement, can be both a dimensionless and a dimensional value (monetary equivalent, quantity of products, etc.).

However, when solving practical problems, situations arise when the effetivity functional of the pair $(\theta_i, \omega_j)$ is a vector and not a scalar. In addition, the exogenous parameters of the effectivity function can be specified as categorical characteristics in case of uncertainty or lack of input information.

In this study, the multicriteria assignment problem is considered, provided that in addition to the cost $c_{ij}$ of assignment $j$-th employee for a certain operation $\theta_i$, a concept of categorical vector that measures effectivity $E_{ij}$ to perform $i$-th operation by $j$-th employee is proposed.

The object of the study is the process of solving a multi-criteria assignment problem in conditions of lack of input information.

The subject of the study is an optimization method for solving a multi-criteria assignment problem with categorical exogenous parameters.

The purpose of this work is to build a mathematical model and, on this basis, to carry out a numerical study of the optimization assignment problem taking into account matrix representation of the effectivity criterion to select optimal pairs $(\theta_i, \omega_j)$ on the Cartesian product of sets $\Theta \times \Xi$.

To achieve this goal, the following tasks have been solved further:

– to develop and substantiate the methodology of numerical coding of categorical data;

– to determine a means of reducing a multicriteria problem to a set of scalar optimization problems;

– to carry out the software implementation of the model based on the creation of a software simulator, and to build the Pareto set of the assignment problem;

– to justify the possibility of generalizing the proposed tools to other classes of assignment problems.

We chose Python as the working programming language, which, thanks to the availability of numerous scientific libraries and cloud computing tools, allows to create the necessary tools for the implementation of numerical experiments.

## 1 PROBLEM STATEMENT

Let for each operation $\theta_i$ of the set $\Theta = \{\theta_i\}$, $i$=1,...,$I$, be given an ordered set

$$v_i = (v_{i1}, v_{i2},\ldots,v_{ik},\ldots, v_{iK}),\ i=1,...,I,$$

containing assessments of operation characteristics and each employee $\omega_j$ of team $\Xi = \{\omega_j\}$, $j$=1,...,$J$ owns an ordered set of values

$$s_j = (s_{j1}, s_{j2},\ldots, s_{jk},\ldots, s_{jK}),\ j=1,...,J,$$

of competence levels, respectively.

Moreover, each value $v_{ik}$ of $i$-th operation characteristics $v_i$ as well as value $s_{jk}$ of $j$-th employee competence level $s_j$ is defined as a categorical parameter.

For each pair $(\theta_i, \omega_j)$, defined on the Cartesian product of sets $\Theta \times \Xi$, the cost $c_{ij}$ of performing an


DOI 10.15588/1607-3274-2024-4-7

operation $\theta_i$ by an employer $\omega_j$ has been set. So, as input data, we also have a matrix of values

$$C = ||c_{ij}||_{i=1,...,I,\ j=1,...,J}.$$

As the endogenous variables of the assignment problem under consideration, we will take the binary matrix $X = ||x_{ij}||_{i=1,...,I,\ i=1,...,J}$, where $x_{ij}$ is the sign of the assignment: $x_{ij} = 1$ of $j$-th employee for the $i$-th operation and $x_{ij} = 0$ otherwise.

Then the assignment problem in the two-criteria statement is to define the optimal values of two performance criteria, namely the cost of execution $G(C,X)$ and the effectivity of execution $\Phi(E,X)$ of operations:

$$\{ G(C,X) \to \min,\ \Phi(E,X) \to \max \} \qquad (1)$$

subject to

$$\sum_{j=1}^{J} x_{ij} \le 1, i = 1,..., I;$$

$$\sum_{i=1}^{I} x_{ij} = 1,\ j = 1,..., J; \qquad (2)$$

$$x_{ij} \in \{0,1\},\ i=1,...,I,\ j=1,...,J.$$

The cost $G(C,X)$ of performing operations of the set $\Theta$ is given by the formula

$$G(E.X) = \sum_{i=1}^{I} \sum_{j=1}^{J} c_{ij} x_{ij}\ .$$

The efficiency $\Phi(E,X)$ is given as follows

$$\Phi(E, X) = \sum_{i=1}^{I} \sum_{j=1}^{J} e_{ij} x_{ij},$$

where the elements $e_{ij}(v_i, s_j)$ of the effectivity matrix $E$

$$E = || e_{ij}(v_i, s_j) ||_{i=1,...,I,\ j=1,...,J}$$

are functional values defining similarity of the vectors ($v_i$, $s_j$).

In other words, functional $e_{ij}(v_i, s_i)$ represents numerical measure of how well $j$-th employee fits to $i$-th operation.

## 2 REVIEW OF THE LITERATURE

Various approaches to solve both deterministic assignment problems and problems under conditions of uncertainty in scalar and multicriteria formulations are proposed by researchers in [1–5] and other studies. At the same time, not only the namely assignment problems have been considered, but also derived classes of problems, for example, the linear problem of ranking alternatives [2]. The total cost and quality of work, the number of performers and the time of work are partial objective functions. An interesting result is the formulation of the assignment problem in fuzzy set terms [1].

E. Acar and H. S. Aplak [4] examine the problem of dynamic distribution of resources for the service provision on the set of locations in urban areas. Service requests occur spontaneously, resources are distributed dynamically. Real-world examples of such applications include dispatching traffic police officers to accident sites and dispatching mechanics to repair sites. It is proposed to use the so-called policy gradient approach, according to which performers are distributed by location in order to minimize the delay.

It is generally not possible to describe exogenous parameters of the mathematical model in numerical form for practical assignment problems. Therefore, certain characteristics of the research objects are presented in a categorical form. Sets of values containing a limited number of separate categories are called categorical (qualitative) ones, while the possible values are parameter levels as it proposed by Kondruk N. E. [6]. At the same time, there is a problem of developing effective distance metrics and measures of similarity for such features, in particular categorical ordered features.

When choosing an approach to the digitization of categorical data from the set of assignment pronlem (1) – (2) input data, a comparative analysis of available classical methods of numerical coding [7] was carried out with the following results:

1. Uniform coding (One-Hot coding) is a process of converting categorical variables into a numerical representation in a binary format, where each category is represented by a binary vector.

Applying the concept of uniform coding to a vector of categorical variables $v_i = (v_{i1}, v_{i2}, v_{iK})$ generates a binary matrix $B^i = \{b^i_{kn}\}_{k=1,...,K,n=1,...,3}$, where the $k$-th row of the matrix $B^i$ corresponds to the categorical variable $v_{ik}$ and contains only one unit, that is, it satisfies the condition

$$\sum_{i=1}^{I} b^i_{kn} = 1, b^i_{kn} \in \{0,1\}.$$

2. Dummy Encoding is a slight improvement over uniform encoding, which uses $N$–1 functions to represent $N$ labels/categories.

3. Ordinal Encoding is used when the categories in the variable have a natural order. In this method, categories are assigned a numerical value based on their order, such as 1, 2, 3, etc.

4. Binary Encoding is similar to One-Hot encoding, but instead of creating a separate column for each category, the latter is represented as binary number.

5. Count Encoding is a method of coding categorical variables by counting the frequency of occurrence of a category in a data set.

6. Target Encoding is a more advanced encoding method used to work with categorical features of high


DOI 10.15588/1607-3274-2024-4-7

power, that is, features with many unique categories. An average target value for each category is calculated, which in turn is used to replace the categorical feature.

Given the conditions of Proposition 1, the most adequate for the case of ordered categorical variables among those mentioned are ordinal or target coding methods.

The Python implementation of these methods, allowing encoding categorical data, is performed using the LabelEncoder and OneHotEncoder encoders from SciKit Learn scientific library [8].

Solving multicriteria assignment problems does not guarantee the generation of a single, best solution, therefore it is important to create and use software tools for constructing a set of Pareto-optimal solutions [9, 10] of problems of this class and conducting numerical experiments.

Currently, such tools are offered as part of a wide range of software products – so-called solvers, i.e. hardware and software for solving complex optimization problems. According to the set of conditions, solvers go through the solution options and choose the best one. They are convenient modern means for creating an information environment to model and solve assignment problems in various fields: in logistics when building the optimal delivery routes and calculating the optimal loading of vehicles; in the retail, medical service, hotel business when generating employee work shift schedules; in industry when solving the ussies of optimizing product warehouses, taking into account the available raw materials.

The market of software products includes commercial developments by IBM or Gurobi, as well as Open Source solvers: OR-Tools, COIN-OR, SCIP, GLPK and others [11–14]. Among the latter, the most developed is the Google OR-Tools solver [11], which is used in the future when creating a software simulator.

## 3 MATERIALS AND METHODS

The statement of the multi-riteria assignment problem (1)–(2) contains numerical (in particular, binary) and categorical types of parameters.

Therefore, at the first stage of processing the mathematical model (1) – (2), there is a need to unify its exogenous parameters.

Let us assume that the characteristics of a certain operation $\theta_i$ constitute a vector of ordered categorical multi-positional variables $v_i = (v_{i1}, v_{i2},\ldots,v_{ik},\ldots, v_{iK})$, where each variable $v_{ik}$ takes a value on a set of M categorical values

$$M = \{\text{“low”, “medium”, “high”}\}.$$

Similar representation has been introduced for vectors $s_i$.

Taking into account the structure of categorical parameters here we propose a novel approach to the digitization of categorical data different from the well-known ones discussed in the previous section.

The approarch involves constructing a function $f(M)$ of the form:

$$M \xrightarrow{f} [0,1], \tag{3}$$

defined on the set $M$. The range of the function $f(M)$ is a unit segment.

A method of digitalization of categorical parameters being proposed in this study is based on the implementation of the well-known Pareto principle [15] and functional-cost analysis [16]. The Pareto principle is one of the universal principles of optimization, whose informal definition is as follows: “20% of the effort provides the opportunity to obtain 80% of the result, and the rest – 80% of the effort – gives only 20% of the result” [17]. In general, an 80/20 ratio often occurs, but does not necessarily have to be exactly 80/20. It can be 90/10, 70/30, etc., which also confirms the observation formulated above.

First of all, we will apply interval estimates of categorical variables, considering the coding problem as continuous. Categorical values must then be encoded by numerical half-open intervals of the form

$$\begin{array}{l} M = \\ \{\text{“low”, “medium”, “high”}\}. \\ \quad\downarrow \qquad\quad \downarrow \qquad\quad \downarrow \\ \{\ [0, N_1), \quad [N_1, N_2), \quad [N_2, 1]\ \} \end{array} \tag{4}$$

We apply the Pareto principle to determine the distribution of the employees’ competence levels and construct the function $f(M)$ of the form (3).

Proposition 1. Based on the Pareto principle, we consider the following distribution of the employees’ competence levels:

$$[0, 0.2), [0.2, 0.8), [0.8,1]. \tag{5}$$

Thus, we will assume (Table 1) that the «high» level of a competence $s_{ik}$ is estimated by the segment [0.8, 1] and means the immediate readiness of proven successful application of knowledge and skills in practice.

The “medium” level of competence means the readiness to apply existing knowledge and skills along with the need for additional training that does not have a critical impact on the quality and time of operation performance, and finally, the “low” level designates the absence or insufficient availability of knowledge that requires additional specialized training.

Table 1 – Interval coding of ordered categorical data

| № of category | Categorical value | Numerical value |
|---|---|---|
| 1 | low | 0… < 0.2 |
| 2 | medium | 0.2...< 0.8 |
| 3 | high | 0.8...1 |


DOI 10.15588/1607-3274-2024-4-7

Proposition 2. The distribution (3) for the digitization of the characteristics $v_{ik}$ which are inherent to $i$-th operation from the set Θ can be explained as follows: the "high" level of the characteristic $v_{ik}$ as a resource of the operation is estimated by the segment [0.8, 1] and means the critical importance of the presence of this resource, the "medium" level means an optional requirement that does not have a critical impact on the execution time of the operation with acceptable quality and, finally, "low" level makes it clear that the presence of this resource is desirable to support the quality assurance of the operation.

In favor of this hypothesis there is the ideology of functional-cost analysis [16]. Within the framework of this paradigm we know that a qualitative dependence of effectivity $K_E$ on cost $K_C$ is described the same way for many technical, economic, and social processes.

Cost $K_C$ can be perceived, for example, as the time or cost of retraining and upgrading the skills of employeers.

Such a dependence is well described by a logistic ($S$-shaped) curve of a kind (Fig. 1):

$$K_E = \frac{1}{1+\exp(-K_C)}. \qquad (6)$$

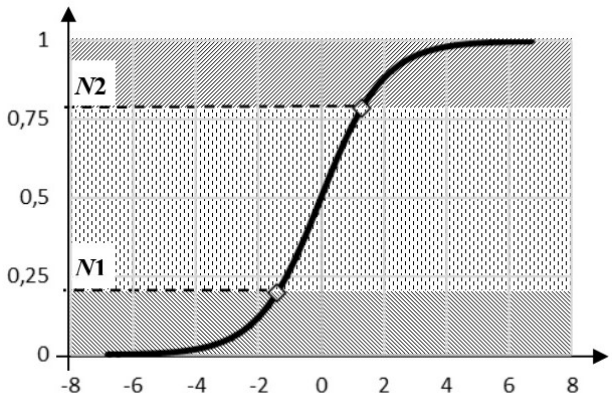



Figure 1 – Decomposition of the logistic curve "effectivity $K_E$ – cost $K_C$"

Fig. 1 shows a graphical presentation of dependence (6) (to within its own affine transformation) together with a horizontal decomposition of the logistic curve into three segments belonging to

– Zone I : [0, $N_1$],
– Zone II : [$N_1$, $N_2$ ],
– Zone III : [$N_2$, 1], $N_1 < N_2$.

Table 1 contains the results concerning approximation of the logistic curve parts belonging to Zones I – III.

Table 1 – Approximation of logistic curve

| № of Zone | Type of approximation dependence $\hat{K}_E$ | $R^2$ |
|---|---|---|
| I | $0.0074\hat{K}_c^2 - 0.0555\hat{K}_c + 0.093$ | 0.934 |
| II | $0.2261\,K_C - 1.76$ | 0.998 |
| III | $0.0087\hat{K}_c^2 - 0.2768\hat{K}_c + 1.201$ | 0.943 |

Thus, estimates of the reliability $R^2$ of the approximating dependence $\hat{K}_E$ confirm the use of such a decomposition. At the same time, horizontal boundary lines of the zones I–III (Fig. 1) practically coincide with the boundaries of the coding ranges (5) of ordered categorical data: $N_1 \approx 0.198$, $N_2 \approx 0.785$.

In this connection, it is interesting to turn to the analysis of the second derivative of the logistic function (6), which has the form:

$$\frac{d^2K_E}{dK_c{}^2} = \frac{2\exp(-2K_c) - \exp(-K_c)}{(1+\exp(-K_c))^3}.$$

On the graph of the second derivative of the logistic function (Fig. 2), two extrema can be distinguished, the arguments of which coincide with the border values $N_1$ and $N_2$.

So, in terms of the main research problem, it can be interpreted as follows: the improvement of the employee's qualification, that is, the improvement of his efficiency, within the limits of a low level, occurs quickly at insignificant costs. We can see that the second derivative is positive and increases, that is, the growth rate of function increases.

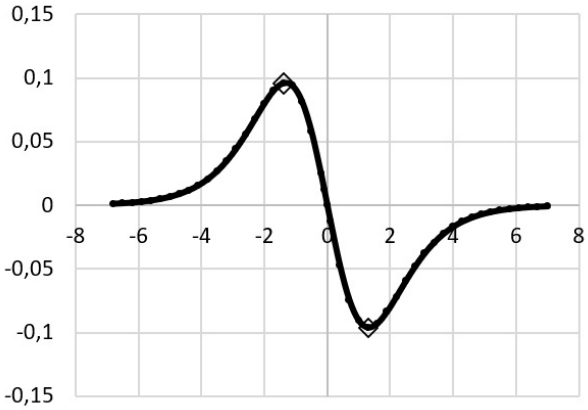



Figure 2 – Second derivative of the logistic curve "effectivity $K_E$ – cost $K_C$"

Within the average level, the second derivative linearly decreases, so that the growth rate of the efficiency function decreases.

This means that more costs are required to improve efficiency within the medium range.


DOI 10.15588/1607-3274-2024-4-7

In the zone of a high level, the second derivative of the logistic function (6) from negative values approaches 0, that is, the rate of change of the function becomes constant.

The type of approximation dependence (Table 1) is convenient to use in the case of the need to implement continuous numerical coding of categorical data.

Let us investigate a scalarization procedure for of partial efficiency criteria, being considered.

In the problem under consideration, the functional $\Phi(E, X)$ of operational efficiency is dimensionless, while the cost $G(C,X)$ of operations according to the statement is a dimensional parameter. Therefore, when choosing a scalarization of partial criteria further, we will normalize the cost matrix $C$ and obtain a normalized matrix $C^H = \| c_{ij}^{norm} \|$ according to the formula

$$c_{ij}^{norm} = \frac{c_{ij} - c^{\min}}{c^{\max} - c^{\min}}.$$

Carrying out numerical coding of categorical characteristics $v_i = (v_{i1}, v_{i2},\ldots,v_{ik},\ldots, v_{iK})$ of operations and levels of competences $s_i = (s_{i1}, s_{i2},\ldots, s_{ik},\ldots, s_{iK})$ of employers according to the scale of Table 1 makes it possible to consider the formalization of the following observation. In practice, certain properties of some operation are more important for its development than others, regardless of the level of qualification features or, in general, the quantity and quality of the necessary resources.

Proposition 3. Let us assume that the characteristics of each $i$-th operation make up the following vector $v_i = (v_{i1}, v_{i2}, v_{iK})$, which assumes a partial or complete ordering by its importance:

$$v_i \xrightarrow{\alpha_i} w_i,$$

where the arguments of the vector $w_i$ are the components of the vector $v_i$, weighted by the components of the vector of preferences

$$\alpha_i=\{\alpha_{i1}, \alpha_{i2},\ldots \alpha_{iK}\}.$$

As an initial approximation for each $i$-th operation, a vector of preferences $\alpha_i=\{\alpha_{i1}, \alpha_{i2},\ldots \alpha_{iK}\}$. can be considered, which specifies the ordering of the operation's characteristics by importance and is calculated according to the formulas

$$\alpha_{ik} \in \left\{ m\left(\sum_{k=1}^{K} k\right)^{-1} \right\}_{m=1,2,\ldots,K}, \sum_{k=1}^{K} \alpha_{ik} = 1, i = 1,2,\ldots, I.$$

In total, the vectors of preferences form a matrix $A=\{\alpha_{ik}\}_{i=1,2,\ldots,I,\ k=1,2,\ldots,K}$ of weighting factors for a set of operations $\Theta$. Accordingly, the vectors $s_j = (s_{j1}, s_{j2}, \ldots, s_{jK})$ of the employers' competencies are also ordered using the matrix $A$, forming the vectors $r_i = (\alpha_{i1}r_{i1}, \alpha_{i2}r_{i2},\ldots \alpha_{iK}r_{iK}).$.

Convolution of this information can be done by several means. This study develops an approach in which the efficiency $e_{ij}$ is estimated by the value of the cosine of the angle $\phi$ between $K$-dimensional vectors $v_i$ and $s_i$ using the scalar product $<v_i, s_j>$ of these vectors of the form:

$$e_{ij} = \cos\phi_{ij} = \frac{< v_i, s_j >}{| v_i | | s_j |}.$$

Thus, as a result of the performed transformations, it is possible to apply a scalarization of the partial criteria $G(C,X)$ and $\Phi(E,X)$.

In so doing we find that the mathematical model of the multicriteria assignment problem being considered takes the form

$$\max\left( \lambda_1 \sum_{i=1}^{I}\sum_{j=1}^{J} e_{ij}x_{ij} - \lambda_2 \sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij}^{norm} x_{ij} \right) \qquad (7)$$

subject to restrictions (2) and additional conditions regarding weighting factors $\lambda_1$, $\lambda_2$:

$$\lambda_1, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1.$$

## 4 EXPERIMENTS

To conduct numerical experiments based on the obtained mathematical model of the multicriteria assignment problem, such software development tools as Google Colab [18] and Google OR-tools [11] were involved in the work (Fig. 3).

```
# Importing necessary libraries
# Google OR-tools and Python
from ortools.linear_solver import
pywraplp
import numpy as np
import matplotlib.pyplot as plt
import random
from numpy.linalg import norm
```

Figure 3 – Application of Google OR-Tools libraries and scientific Python libraries for modeling

Google OR-Tools is a set of open source software (Apache 2.0 license) developed by Google for solving optimization problems, including mathematical programming, vehicle routing, assignment problems, and others. Google OR-Tools contains a set of ready-made components written in such well-known programming languages as C++, Java, .NET and Python

The assignment problem was solved using the MIPSolver shell, which is a component of OR-Tools,


DOI 10.15588/1607-3274-2024-4-7

designed for linear programming and mixed integer programming (MIP) problems. MIPSolver interacts with the SCIP tool, a platform for solving integer problems with constraints (CIP) and mixed integer nonlinear problems (Fig. 4)

```
# Creating MIP solver, that calls SCIP
# backend
solver =
pywraplp.Solver.CreateSolver("SCIP")
```

Figure 4 – Call a MIPSolver

The stages of numerical modeling and solving the assignment problem in the Google Colab cloud environment using OR-Tools are presented in Fig. 5.

The experiments were conducted on a set of input data {*I*, *J*, *K*, *C*, *S*} (Block 1 in Fig. 5).

At the same time, at the choice of the DM, the parameters of the dimensionality of the problem and the matrices of efficiency *E* and cost *C* can be randomly set by calling the Python functions random.randint and random.random() or read from a data file.

Blocks 2 and 3 of the information model being described are preparatory for calling required functions from OR-Tools arsenal.

Blocks 4–6 serve to describe problem independent (endogenous) variables, constraints, the specification of linear scalarization of two partial criteria of the original problem: efficiency criteria and the cost ones, after that and the transfer of this information is carried out to MIPSolver. At the same time, we consider that the relation $I \geq J$ is fulfilled.

Further, the function solver.Solve() call (Block 7) implements the process of solving two-criterion assignment problem with the given values of the weighting coefficients $\lambda_1$ and $\lambda_2$.

## 5 RESULTS

The last stage of modeling (Block 8) is intended to construct and analyse a set of Pareto-optimal solutions of the two-criteria assignment problem under consideration.

The set of Pareto-optimal solutions, which is the result of solving the two-criteria assignment problem for definite set of exogenous parameters {*I*, *J*, *K*, *C*, *S*}, is constructed in the space of partial criteria based on the variation of their weighting coefficients $\lambda_1$ and $\lambda_2$ accordingly.

Thus, the procedure for numerical modeling of the assignment problem necessarily involves the visualization of a set of Pareto-optimal solutions in the space of partial target criteria, an example of which is shown in Fig. 6.

```
# Setting input data -                          Block 1
# exogenous parameters:
num_workers #J
num_tasks #I
num_charact #K
efficiency_workers #S[][]
charact_tasks #V[][]
cost_workers #C[][]
```

```
# Standardization of the                        Block 2
# cost matrix: cost_workers
```

```
# Digitization of exogenous                     Block 3
# categorical parameters:y_dot[][]
```

```
# Problem endogenous variables                  Block 4
x[i,j] = solver.IntVar(0, 1, "")
```

```
# Problem constraints                           Block 5
solver.Add(solver.Sum([x[i, j]
for j in range(num_tasks)]) <= 1)
solver.Add(solver.Sum([x[i, j]
for i in range(num_workers)]) == 1)
```

```
# Objective functional                          Block 6
# Weight1: λ1, Weight1: λ2
objective_terms.append((Weight1*
y_dot[][] - Weight2*cost_workers[][])*
x[][])
```

```
# Solving the problem                           Block 7
status = solver.Solve()
```

```
# Analysis of Pareto-optimal                    Block 8
# solutions
# Definining "indifference" ranges
# of partial objective criteria
```

Figure 5 – Numerical modeling of the assignment problem in the Google OR-Tools environment

The analysis of a set of Pareto-optimal solutions enables the DM to make flexible decisions regarding the final choice.

Numerous numerical experiments with various data sets provide the basis for the following constructive remarks.

Remark 1. Values of partial effectivity and cost criteria are unchanged (indifferent) on certain ranges of values of weight coefficients $\lambda_1$, $\lambda_2$, despite different values of the generalized criterion (7).


DOI 10.15588/1607-3274-2024-4-7

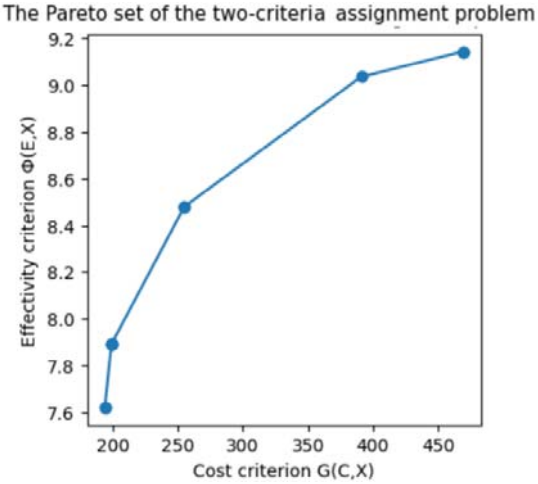



Figure 6 – Construction of the Pareto set for one of sets {*I*, *J*, *K*, *C*, *S*} of exogenous parameters

This result is statistically significant, because it was observed in a critically large part of the numerical experiments. An illustration of this conclusion is shown in the Fig. 6, which provides an example of solving a certain assignment problem, where the ranges for determining the weighting factors $\lambda_1$, $\lambda_2$, contained 10 values each. And only five pairs of them generated different results. In other words, sets of $\lambda_1$, $\lambda_2$ values {(0;1), (0.1;0.9)}, {(0.2;0.8)}, {(0.3;0.7),(0.4;0.6)}, {(0.5;0.5), (0.6;0.4), (0.7;0.3), (0.8;0.2)} {(0.9;0.1), (1;0)}, enclosed in curly braces, generate the same results respectively. In Fig. 6 different values pairs of partial criteria are shown by dots (markers) – generated different results. The solid line in Fig.6 represents an approximate shape of the Pareto set.

Remark 2. Further processing of information regarding the set of Pareto-optimal solutions obtained allows to define set of ranges, which are characterized by relative indifference of partial criteria values. In so doing the DM can apply of a certain concession on the values of the partial criteria which he considers as acceptable.

## 6 DISCUSSIONS

The proposed mathematical model of the two-criterion assignment problem, unlike the classical formulation, takes into account the possibility of specifying exogenous parameters with categorical values, which expands the range of model applications. Introducing logistic mapping into consideration at the stage of digitization of categorical variables, in contrast to known methods, the analysis of which is carried out in the section “Review of the literature”, allows for finer adjustment to the specifics of the problem and, if necessary, to apply elements of continuous analysis.

The proposed algorithm and software form a simple flexible and reproducible environment for the numerical study of multi-criteria assignment problems. The presented calculation results confirm the validity and efficiency of the developed mathematical model and solution method for multi-criteria assignment problem being considered.

## 7 CONCLUSIONS

Modeling and numerical research of the multi-criteria assignment problem was carried out under the condition of vector presentation of the efficiency criterion and the presence of categorical exogenous parameters.

The scientific novelty.

We proposed a novel method of solving the assignment problem that implemented as a multi-stage process, which includes the stage of transformation of exogenous parameters of the model, given by categorical variables, based on the implementation of the Pareto principle and logistic mapping, the stage of constructing linear scalarization of the efficiency and the cost criteria.

A generalization of the multi-criteria assignment problem has been developed, where the partial criterion for the performance of *i*-th operation by *j*-th employee is defined as the scalar product of the digitized competences of the employee and the characteristics of the operation.

The practical significance.

We applied cloud software development tools of Google Colab and Google OR TOOLS as well as Python scientific libraries to conduct numerical experiments on the basis of the mathematical model proposed for the main research problem.

Numerical experiments conducted on multiple sets of input data using this optimization mathematical model showed that the proposed constructive means of mathematical modeling and information environment developed in this study provide flexible methodology for researching other types of optimization assignment problems.

Prospects for further research consist in developing the proposed tools in two directions. First, it is an adaptation of the proposed model by considering exogenous parameters of assignmet efectivity as functions of time and/or cost. The second direction of research is in developing the solution method to reach refined approximation of the set of Pareto-optimal solutions for considered multi-criteria assignment problem.

## ACKNOWLEDGMENTS

The study was performed within the framework of the state-budget research “City’s information infrastructure modeling and optimizatione” of O.M. Beketov National University of Urban Economy in Kharkiv (state registration number 0121U109215).

## REFERENCES

1. Song M., Cheng L. Solving the reliability-oriented generalized assignment problem by Lagrangian relaxation and Alternating Direction Method of Multipliers, *Expert Systems with Applications*, 2022, Vol. 205, P. 117644. DOI: 10.1016/j.eswa.2022.117644
2. Mehlawat M. K., Gupta P., Pedrycz W. A New Possibilistic Optimization Model for Multiple Criteria Assignment Problem, *IEEE Transactions on Fuzzy Systems*, 2018, Vol. 26, № 4, pp. 1775–1788. DOI: 10.1109/TFUZZ.2017.2751006
3. Borovicka A. Possible Modifications of the Multiple Criteria Assignment Method, *Czech Economic Review*, 2013, Vol. 7, pp. 55–67.


DOI 10.15588/1607-3274-2024-4-7

4. Acar E., Aplak H. S. A Model Proposal for a Multi-Objective and Multi-Criteria Vehicle Assignment Problem: An Application for a Security Organization, *Mathematical and Computational Applications*, 2016, Vol. 21(4), P. 39. DOI: 10.3390/mca21040039
5. Yan Y., Deng Y., Cui S. et al. A policy gradient approach to solving dynamic assignment problem for on-site service delivery, *Transportation Research Part E: Logistics and Transportation Review*, 2023, Vol. 178, P. 103260 DOI: 10.1016/j.tre.2023.103260
6. Kondruk N. E. Methods for determining similarity of categorical ordered data, *Radio Electronics, Computer Science, Control*, 2023, № 2, pp. 31–36. DOI: 10.15588/1607-3274-2023-2-4.
7. Dahouda M. K., Joe I. A Deep-Learned Embedding Technique for Categorical Features Encoding, *IEEE Access*, 2021, Vol. 9, pp. 114381–114391. DOI: 10.1109/ACCESS.2021.3104357.
8. Scikit-learn 1.4.2 [Electronic resource]. Access mode: https://pypi.org/project/scikit-learn/
9. Dosantos P. S., Bouchet A., Mariñas-Collado I. et al. OPSBC: A method to sort Pareto-optimal sets of solutions in multi-objective problems, *Expert Systems With Applications*, 2024, Vol. 250, P. 123803. DOI: 10.1016ps://doi.org//j.eswa.2024.123803.
10. Makarova V.V. Pryntsyp Pareto v konteksti orhanizatsii staloho ahrarnoho zemlekorystuvannia, *Pryazovskyi ekonomichnyi visnyk*, 2020. Vyp. 1(18), pp. 220–224. DOI: https://doi.org/10.32840/2522-4263/2020-1-39
11. Route. Schedule. Plan. Assign. Pack. Solve. OR-Tools is fast and portable software for combinatorial optimization [Electronic resource]. Access mode: https://developers.google.com/optimization.
12. COIN-OR [Electronic resource]. Access mode: https://www.coin-or.org/
13. SCIP. Solving Constraint Integer Programs. [Electronic resource]. – Access mode: https://www.scipopt.org/
14. GLPK (GNU Linear Programming Kit) [Electronic resource]. – Access mode: https://www.gnu.org/software/glpk/
15. Koh R. Princip 80/20 ta 92 fundamental'nih zakoni prirodi. Nauka uspihu. Kiyv, Vidavnichij centr "KM-BUKS", 2019, 360 p.
16. Petrov E. H., Novozhylova M. V., Hrebennik I. V. Metody i zasoby pryiniattia rishen u sotsialno-ekonomichnykh systemakh. Kyiv, Tekhnika, 2003, 240 p.
17. Blackman I., Chan E. Pareto principle plus statistic methodology in establishing a cost-estimating model, *International CIB World Building: 19th Congress, Brisbane, 05–09 May 2013: proceedings*. Brisbane, Queensland University of Technology, 2013, pp. 1–14.
18. Google-Colab [Electronic resource]. Access mode:https://colab.research.google.com/?utm_source=scs-index.

Received 15.05.2024.
Accepted 09.10.2024.

УДК 519.85

## РОЗВ'ЯЗАННЯ БАГАТОКРИТЕРІАЛЬНОЇ ЗАДАЧІ ПРО ПРИЗНАЧЕННЯ ЗА УМОВИ КАТЕГОРІАЛЬНОГО КРИТЕРІЮ ЕФЕКТИВНОСТІ

**Новожилова М. В.** – д-р фіз.-мат. наук, професор, завідувач кафедри комп'ютерних наук та інформаційних технологій, Харківський національний університет міського господарства імені О. М. Бекетова, Україна.

**Карпенко М. Ю.** – канд. техн. наук, доцент, доцент кафедри комп'ютерних наук та інформаційних технологій, Харківський національний університет міського господарства імені О. М. Бекетова, Україна.

### АНОТАЦІЯ

**Актуальність**. В роботі розглядається задача про призначення множини виконавців на скінчений набір операцій в багатокритеріальній постановці за умови ієрархічної структури частинного критерію ефективності виконання набору операцій, поданого таким чином, що кожний виконавець володіє скінченим набором компетентностей і кожна операція має скінчений набір характеристик. В якості екзогенних параметрів задачі передбачено застосування числових та категоріальних типів даних. Актуальність задачі про призначення, що розглядається, обумовлюється надзвичайно широким спектром практичних застосувань як в класичній постановці, так і нових модифікаціях, запит на які постійно генерується економічним середовищем, що динамічно розвивається. При цьому критично менша кількість наукових публікацій присвячена засобам моделювання та розв'язання саме багатокритеріальних задач про призначення, незважаючи на важливість задач такого типу в прийнятті рішень, як в теоретичному, так і в практичному аспектах. В загальному випадку в умовах нестачі інформації екзогенні параметри задачі не можуть бути задані в числовій формі, тому виникає потреба застосування категоріальних даних з подальшим числовим кодуванням.

**Метою** роботи є побудова математичної моделі та проведення на цій основі чисельного дослідження оптимізаційної задачі про призначення з урахуванням можливості ієрархічної структури частинного критерію ефективності вибору пар «операція – виконавець».

**Метод** розв'язання задачі – це багатоетапний процес, що включає етап трансформації екзогенних параметрів моделі, заданих категоріальними змінними, на основі втілення принципу Парето та логістичного відображення, етап побудови лінійної згортки частинних критеріїв ефективності та вартості виконання робіт.

**Результати роботи.** Проведено системологічний аналіз існуючих підходів та запропоновано методику числового кодування категоріальних даних. Обґрунтовано засіб зведення багатокритеріальної за постановкою задачі до набору однокритеріальних оптимізаційних задач на основі лінійної згортки та функціонально-вартісного аналізу, побудовано множину Парето основної задачі дослідження.

**Висновки**. Проведене моделювання та розв'язання задачі про призначення на основі створення програмного симулятора із застосуванням солверу Google OR-Tools підтвердило можливість узагальнення запропонованих інструментальних засобів на інші класи задач про призначення.

**КЛЮЧОВІ СЛОВА**: математичне та комп'ютерне моделювання задачі про призначення, багатокритеріальна оптимізація, множина Парето, категоріальні параметри, логістична крива.


DOI 10.15588/1607-3274-2024-4-7

**ЛІТЕРАТУРА**

1. Song M., Solving the reliability-oriented generalized assignment problem by Lagrangian relaxation and Alternating Direction Method of Multipliers / M. Song, L. Cheng // Expert Systems with Applications. – 2022. – Vol. 205. – P. 117644. DOI: 10.1016/j.eswa.2022.117644
2. Mehlawat M. K. A New Possibilistic Optimization Model for Multiple Criteria Assignment Problem / M. K. Mehlawat, P. Gupta, W. Pedrycz // IEEE Transactions on Fuzzy Systems. – 2018. – Vol. 26, №4. – P. 1775–1788. DOI: 10.1109/TFUZZ.2017.2751006
3. Borovicka A. Possible Modifications of the Multiple Criteria Assignment Method / A. Borovicka // Czech Economic Review.– 2013. – Vol. 7. – P. 55–67.
4. Acar E. A Model Proposal for a Multi-Objective and Multi-Criteria Vehicle Assignment Problem: An Application for a Security Organization / E. Acar, H. S. Aplak // Mathematical and Computational Applications. – 2016. – Vol. 21(4), P. 39. DOI: 10.3390/mca21040039
5. A policy gradient approach to solving dynamic assignment problem for on-site service delivery / [Y. Yan, Y. Deng, S. Cui et al.] // Transportation Research Part E: Logistics and Transportation Review. – 2023. – Vol. 178, P. 103260. DOI: 10.1016/j.tre.2023.103260
6. Kondruk N. E. Methods for determining similarity of categorical ordered data / N. E. Kondruk // Radio Electronics, Computer Science, Control. – 2023. – № 2. – P. 31–36. DOI: 10.15588/1607-3274-2023-2-4.
7. Dahouda M. K. A Deep-Learned Embedding Technique for Categorical Features Encoding / M. K. Dahouda, I. Joe // IEEE Access. – 2021. – Vol. 9. – P. 114381–114391. DOI: 10.1109/ACCESS.2021.3104357.
8. Scikit-Learn 1.4.2 [Electronic resource]. – Access mode: https://pypi.org/project/scikit-learn/
9. OPSBC: A method to sort Pareto-optimal sets of solutions in multi-objective problems / [P. S. Dosantos, A. Bouchet, I. Mariñas-Collado et al.] // Expert Systems With Applications. – 2024. – Vol. 250. – P. 123803 DOI: https://doi.org/10.1016/j.eswa.2024.123803.
10. Макарова В.В. Принцип Парето в контексті організації сталого аграрного землекористування / В. В. Макарова // Приазовський економічний вісник. – 2020. – Вип. 1(18). – С. 220–224. DOI: https://doi.org/10.32840/2522-4263/2020-1-39.
11. Route. Schedule. Plan. Assign. Pack. Solve. OR-Tools is fast and portable software for combinatorial optimization [Electronic resource]. – Access mode: https://developers.google.com/optimization.
12. COIN-OR [Electronic resource]. – Access mode: https://www.coin-or.org/
13. SCIP. Solving Constraint Integer Programs. [Electronic resource]. – Access mode: https://www.scipopt.org/
14. GLPK (GNU Linear Programming Kit) [Electronic resource]. – Access mode: https://www.gnu.org/software/glpk/
15. Кох Р. Принцип 80/20 та 92 фундаментальних закони природи. Наука успіху / Р. Кох. – Київ : Видавничий центр «КМ-БУКС», 2019. – 360 с.
16. Петров Е. Г. Методи і засоби прийняття рішень у соціально-економічних системах / Е. Г. Петров, М. В. Новожилова, І. В. Гребеннік. – Київ : Техніка, 2003. – 240 с.
17. Blackman I. Pareto principle plus statistic methodology in establishing a cost-estimating model / I. Blackman, E. Chan // International CIB World Building: 19th Congress, Brisbane, 05–09 May 2013: proceedings. – Brisbane : Queensland University of Technology, 2013 – P. 1–14.
18. Google-Colab [Electronic resource]. – Access mode:https://colab.research.google.com/?utm_source=scs-index.


DOI 10.15588/1607-3274-2024-4-7

p-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2024. № 4
e-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2024. № 4

# НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

# NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 004.8

## ENSEMBLE OF SIMPLE SPIKING NEURAL NETWORKS AS A CONCEPT DRIFT DETECTOR

**Bodyanskiy Ye. V.** – Dr. Sc., Professor, Professor of the Department of Artificial Intelligence of the Kharkiv National University of Radioelectronics, Kharkiv, Ukraine.

**Savenkov D. V.** – Postgraduate student of the Department of Artificial Intelligence of the Kharkiv National University of Radioelectronics, Kharkiv, Ukraine.

### ABSTRACT

**Context.** This paper provides a new approach in concept drift detection using an ensemble of simple spiking neural networks. Such approach utilizes an event-based nature and built-in ability to learn spatio-temporal patterns of spiking neurons, while ensemble provides additional robustness and scalability. This can help solve an active problem of limited time and processing resources in tasks of online machine learning, especially in very strict environments like IoT which also benefit in other ways from the use of spiking computations.

**Objective.** The aim of the work is the creation of an ensemble of simple spiking neural networks to act as a concept drift detector in the tasks of online data stream mining.

**Method.** The proposed approach is primary based on the accumulative nature of spiking neural networks, especially Leaky Integrate-and-Fire neurons can be viewed as gated memory units, where membrane time constant $\tau_m$ is a balance constant between remembering and forgetting information. A training algorithm is implemented that utilizes a shallow two-layer SNN, which takes features and labels of the data as an input layer and the second layer consists of a single neuron. This neuron's activation implies that an abrupt drift has occurred. In addition to that, such model is used as a base model within the ensemble to improve robustness, accuracy and scalability.

**Results.** An ensemble of shallow two-layer SNNs was implemented and trained to detect abrupt concept drift in the SEA data stream. The ensemble managed to improve accuracy significantly compared to a base model and achieved competitive results to modern state-of-the-art models.

**Conclusions.** Results showcased the viability of the proposed solution, which not only provides a cheap and competitive solution for resource-restricted environments, but also open doors for further research of SNN's ability to learn spatio-temporal patters in the data streams and other fields.

**KEYWORDS:** machine learning, online learning, spiking neural networks, concept drift, drift detector, artificial neural networks, data stream mining, artificial intelligence, leaky integrate-and-fire neuron.

### ABBREVIATIONS

ANN is an artificial neural network;
GRU is a gated recurrent unit;
LSTM is a long short-term memory;
ML is a machine learning;
PLIF is a parametric leaky integrate-and-fire neuron;
RNN is a recurrent neural network;
SEA is a streaming ensemble algorithm;
SNN is a spiking neural network.

### NOMENCLATURE

$\tau_m$ is a membrane time variable that determines the decay rate;
$f(V_{t-1}, X_t)$ is a LIF's neural dynamics function;
$g(X)$ is an ensemble's function;
$h_j(X)$ is an ensemble's $j$-th base model function;
$N$ is a number of base models used in ensemble;
$T$ is a number of iterations the SNN receives the input;
$V_t$ is a membrane potential at a time point $t$;
$X_t$ is an input vector at a time point $t$;
$x_i$ is a data stream's $i$-th feature;
$y$ is a data stream's true label.
$y'$ is a model's output of drift detection.

### INTRODUCTION

Data stream mining in the online manner is a complex task that involves strict restrictions on time and processing resources, while still requiring adequate results and a need to process huge volumes of data. This implies con-


DOI 10.15588/1607-3274-2024-4-8

siderable limitations on used algorithms, even making many state-of-the-art ML models unusable. In addition, many models are unable to adapt quickly to the changing environment and therefore become obsolete. This is especially prominent with ANNs which fail to adjust their trainable parameters to the concept drift without utilizing resource-hungry recurrent architectures [1].

Many approaches have been proposed to solve existing online learning problems. In recent years, interest has spiked greatly in applying more biologically inspired systems to solve ML tasks. SNNs are one such system. They are considered to be the third generation of ANNs that closely mimic the inner work of biological neurons in human brains [2]. In comparison to regular ANNs, spiking neurons implement biological mechanisms like the accumulation of membrane potential, refractory period, charge decay and simulation of neurotransmitters' explosions. Such mechanisms result in lower power consumption [3], analogue and event-based nature, as well as an ability to learn temporal patterns in the data [4].

These characteristics make SNNs a viable solution for listed problems of online learning, especially in the most resource limited environments like IoT, which also benefit from other advantages of spiking computations. Recent research showcased that SNNs are able to adapt to the concept drift [5] and to be used as drift detectors as well when combined with evolving architectures [6].

But overall, the research on spiking drift detectors is very limited. In addition to that and to our knowledge, no research was proposed to utilize ensembling techniques with SNNs to detect the concept drift, while ensembles of ANNs were showcased to be efficient in such task, as well as improving SNNs performance (as described in section 2).

**The object of study** is the process of concept drift detection during online learning and data stream mining.

**The subject of study** is the use of spiking neural networks as concept drift detectors.

**The purpose of the work** is to broad the existing research by exploring how SNN's built-in abilities to learn spatio-temporal patterns can be utilized to build spiking detectors with shallow architectures and exploring how ensemble of such detectors can create a more robust system.

## 1 PROBLEM STATEMENT

This study aim is to develop an ensemble of simple SNNs to identify the abrupt concept drift in the artificially generated data stream and to analyze whether such approach improves the performance of the base model.

To do so, a training pipeline is to be implemented that trains a model in the online manner by providing the input vector $X$ that consists of both the input features and the true label of the data stream:

$$X = (x_1, x_2, ..., x_n, y), \quad (1)$$

where $x_i$ represents a stream's feature and $y$ is its true label.

As a result, two experiments are to be performed: one with just the SNN model and one with the ensemble; and their metrics are to be compared.

## 2 REVIEW OF THE LITERATURE

In regular ANNs, memorization is achieved by implementing recurrent layers or by using gating mechanisms, like in LSTM [7] or GRU [8]. Previously, researchers tried to reimplement similar memory blocks for SNNs [9], but recently, they focused more on built-in short-term memory capacity. SNNs have an accumulative nature due to the membrane potential and authors of [10] noted that the Leaky Integrate-and-Fire (LIF) model can be viewed as a gated memory unit if the formula is rewritten the next way:

$$f(V_{t-1}, X_t) = V_{t-1} + \frac{1}{\tau_m}(-(V_{t-1} - V_{reset}) + X_t). \quad (2)$$

Which in case of $V_{reset} = 0$ looks like:

$$f(V_{t-1}, X_t) = \left(1 - \frac{1}{\tau_m}\right)V_{t-1} + \frac{1}{\tau_m}X_t. \quad (3)$$

As noted in the original research, the integration process and the leak can be viewed as update and reset gates, while the accumulated membrane potential $V_{t-1}$ functions as a hidden state. With such interpretation, the membrane time constant $\tau_m$ is a balance constant between remembering and forgetting information. A larger $\tau_m$ allows for a better estimate of the frequency of input spikes, but requires more time to charge and fire, which works well with a stable input signal. On the contrary, a smaller $\tau_m$ means a faster voltage decay, which makes the LIF neuron sensitive to time-varying input signals and able to respond quickly – at the cost of losing the accuracy of distinguishing between different input signals. Furthermore, the same research proposed to make the $\tau_m$ trainable to help to find the balance without an additional hyperparameter tuning. Such solution they named as PLIF model.

Therefore, LIF nodes can be used as a cheap memory unit which also have an event-based nature. In addition, LIFs have perks of regular dense layers, like being able to combine inputs and capture relationships between different features. In this context, LIF nodes can function as two layers at the same time. This makes them even more suitable to be used as concept drift detectors, especially when faced with sudden or recurring drifts.

A popular and a common technique in ML is usage of ensemble learning techniques. Ensembles combine predictions of multiple base models which provides an improved generalization, better handling of class imbalance and overall robustness. In addition to that, ensembles can be more resource-efficient than training a single and complex model. Due to these, ensembles gathered attention to


DOI 10.15588/1607-3274-2024-4-8

be used in online learning scenarios and have been already showcased to be effective in them [11, 12], especially when used in IoT systems [13] (where SNNs shine as well [14]). They also showcased the ability to detect the concept drift [15].

The human nervous system also appears to utilize some ensemble effects or techniques: groups of neurons or neural circuits act together to improve motor skills, as was noted in the [16]. The same research also developed and tested Spiking Voter Ensemble Network which was built on ensemble of simplified three-neuron models with utilization of the input timing dependent synaptic plasticity. The proposed model achieved improved MNIST [17] performance. Other research also showcased general improvements for simple SNNs when combined with ensemble techniques: authors of [18] achieved state-of-the-art results for MNIST, NMNIST [19] and DSV Gesture [20] datasets while lowering the number of trainable parameters by half; and authors of [21] also achieved state-of-the-art results for MNIST and CIFAR-10 [22] datasets.

## 3 MATERIALS AND METHODS

The key part of any ensemble algorithm is a base model. As was noted previously, the use of online learning sets restrictions for such models: the model should be simple to not consume large amounts of resources, while also being able to adapt to new patterns in the data stream. In our proposed solution, we utilize a shallow two-layer PLIF-based with the next key characteristics:

1) The first layer takes the input vector $X$ to learn patterns between the features and the labels;

2) The second layer consists of a single PLIF neuron, activation of which indicates if the concept drift occurred;

3) The PLIF's membrane state does not reset between the iterations. This allows the model's internal state to accumulate information over time and capture subtle changes in the data stream that may indicate a potential drift.

The architecture is shown in Fig. 1:

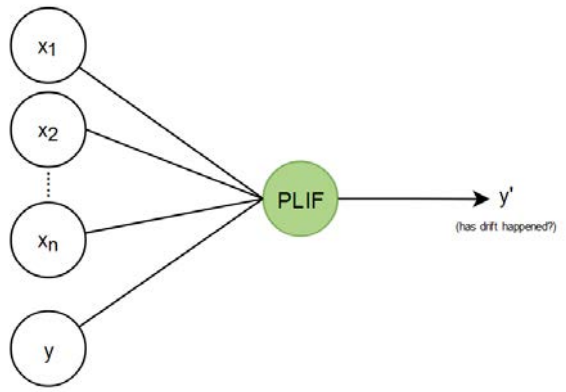



Figure 1 – The proposed model architecture

This approach does not use any additional helping models and only analyzes the stream's data. Ideally, it can learn both virtual and real drifts. In addition, the refractory period, that comes after the action potential, prevents excessive neural activity. Meaning, it won't fire again right after the drift detection. This lowers chances of recurring false-positive activations and removes the need to develop additional timeout mechanisms for the model when it successfully detected the drift. Also, as was mentioned in the section 2, the initial $\tau_m$ controls how sensitive the neuron is. Though it is different for each task, normally it is expected that data stream will be stable and concept drift won't occur often. Because of that, higher values of the initial $\tau_m$ will prevent the fast charge decay and can further decrease false-positive activations.

The main drawback of such SNN model (which will be referred as the base model) comes from its simplicity, meaning it will fail to learn complex patterns without changes in the architecture. To overcome this, an ensemble is used. The ensemble trains up to $N$ base models, each receiving the input vector $X$. The verdict whether concept drift has occured is determined by the formula (4):

$$g(X) = \max(h_1(X), h_2(X), ..., h_N(X)), \quad (4)$$

where $h_j(X)$ represents the base model function. If any of the models are activated, then we conclude that the drift has occurred.

For the model training, a backpropagation algorithm is used. And to ease the experiment, it has a few restrictions on the concept drift: an abrupt concept drift occurs after the set number of iterations and a new drift cannot occur before the previous one was successfully detected.

The whole process is visualized in the Fig. 2.

## 4 EXPERIMENTS

The developed training algorithm was tested using the SEA (streaming ensemble algorithm) artificial data stream generator [23] with an abrupt concept drift. This data stream generates 3 features $\{x_1, x_2, x_3\}$ (where $x_3$ doesn't take part in the classification) and a binary target variable that is calculated by one of the (5–8) equations that switch during the drift:

$$y = (x_1 + x_2) \le 8, \quad (5)$$

$$y = (x_1 + x_2) \le 9, \quad (6)$$

$$y = (x_1 + x_2) \le 7, \quad (7)$$

$$y = (x_1 + x_2) \le 9.5. \quad (8)$$

As was mentioned in section 1, two experiments were performed: one with just the base model and second one with the ensemble of multiple such models. Both experiments run with the next configuration:

– the initial membrane time variable was set at $\tau_m = 5$ for slower charge decay;


DOI 10.15588/1607-3274-2024-4-8

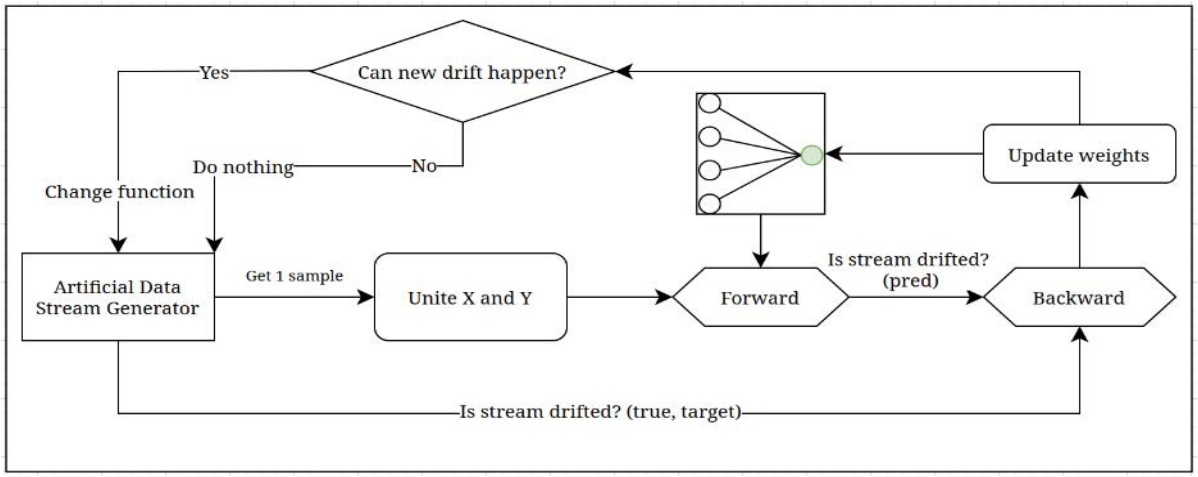



Figure 2 – Flowchart of the training algorithm

– training is done with 20000 iterations, testing is done with 2000 iterations;

– spiking neurons "see" the data only for a $T = 1$ number of iterations;

– a new drift occurs once in 200 iterations.

Initial weights are randomly generated and no additional hyperparameter tuning was performed. The ensemble uses $N = 5$ base models which are trained on the same data. Also, the random seed was freezed.

The tools used for the program realization are Python, PyTorch [24] and SpikingJelly [25].

The results of these experiments are shown in the section 5.

## 5 RESULTS

The first experiment with a simple base model did not yield good results. While this SNN was capable of identifying the drift, the accuracy of predictions stayed very low, with the last accuracy being equal to 29% and the average accuracy stagnating only around 34%. The accuracy history graph is shown in Fig. 3.

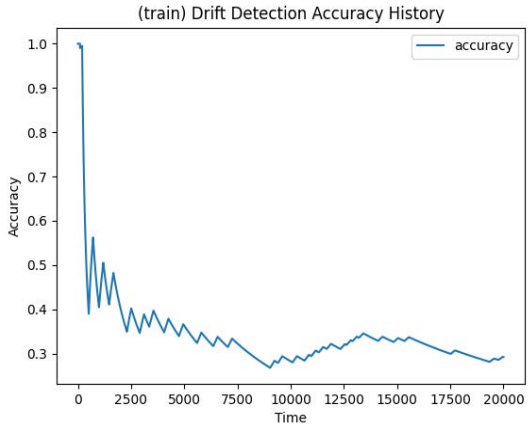



Figure 3 – The training accuracy history of the base spiking model

The situation with the testing is similar with the last accuracy being equal to 43% as shown in Fig. 4 (it also shows when the drift has happened and has been identified). Though, the fact that the metric stayed consistent between the training-testing phases is promising. In addition to that, on average, SNN required 285 iterations before identifying the drift. Also, the overall efficiency is very sensitive to the initial hyperparameters.

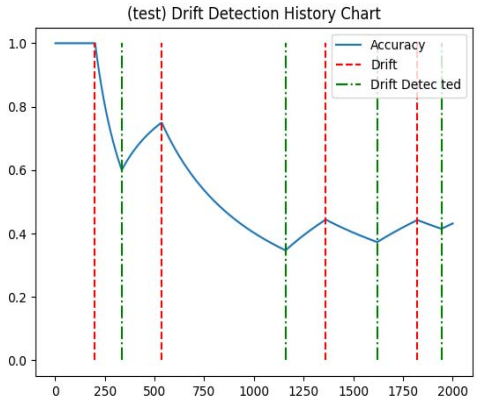



Figure 4 – The testing accuracy history of the base spiking model

The ensemble method managed to produce much better results: the average train accuracy increased to 87% (as shown in Fig. 5) on the train data and maintained the test accuracy around 91% (as shown in Fig. 6). In addition to that, the drift detection time improved as well: on average, the drifts were detected in 22.65 iterations during the training and in 8.44 iteration during the testing.

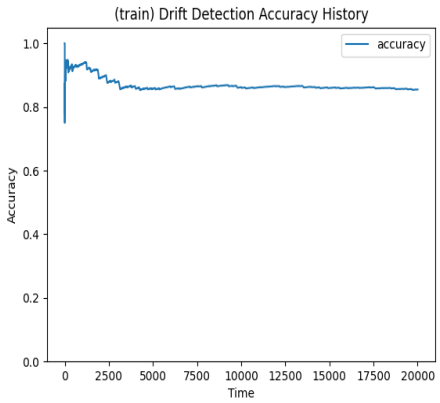



Figure 5 – The training history of the ensemble


DOI 10.15588/1607-3274-2024-4-8

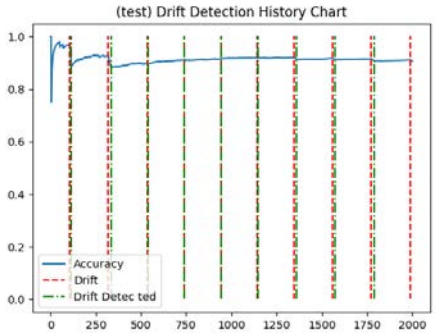



Figure 6 – The testing accuracy history of the ensemble

## 6 DISCUSSION

As evident from the results, the provided base model was very weak and non-robust, making it non-applicable to real-world tasks, especially with more complex data. Still, it showcased the ability to learn the patterns. But when it was used within the ensemble, the overall accuracy and detection time improved significantly.

Moreover, the achieved results are comparable to the current state-of-the-art accuracy which include classical ML algorithms [26] like Naïve Bayes with an 86% accuracy and another ensemble-based Hoeffding Tree classifier which also produced an 86% accuracy score (or 88% when combined with the adaptive sliding window algorithm [11]). Besides, the proposed solution achieved similar results to the current state-of-the-art RNN-based (recurrent neural network) solution [27], which produced an accuracy of 91.70%, in addition to having a much smaller number of trainable parameters.

This indicates the viability of the proposed solution to be used in data streams analysis when faced with resource-restricted environments. Also, such methodology allows for an easier integration into the IoT systems, where each SNN can be implemented as a small separate circuit with analogue sensors as an input. Also, it can function as a plausible alternative to the deep learning with which SNNs often struggle [28].

Nonetheless, additional testing on different data streams is required to improve the proposed architecture and algorithm and to test robustness. As already mentioned, the solution has its own disadvantages: it will struggle with more complex data and will likely require significant changes in the base model architecture.

## CONCLUSIONS

This paper proposed a new approach in developing concept drift detectors that focuses on utilizing SNN's abilities to spot the changes in spatio-temporal patterns in data stream inflicted by the drift, and further improving it with ensembling techniques. Experiments with a common for such tasks artificial data were performed that gave adequate and competitive results, which validated the viability of the approach.

The scientific novelty of this work is that the method of using an ensemble of simple spiking drift detectors is firstly proposed. Ensemble's base models take in features and label to learn spatio-temporal patters and fire if the abrupt drift has occurred; while the ensemble of multiple such models itself significantly improves the overall accuracy and gives state-of-the-art results on the tested artificial data, in addition to having small number of trainable parameters.

The practical significance of obtained results is that such approach provides a cheap and efficient solution for extremely resource-limited environments that face constant dangers of concept drift occurrences, like many IoT systems that also benefit from analog and event-based nature of SNNs. Moreover, the use of ensembles provides an additional scalability by allowing the increase or decrease of model quantity to fit the practical needs.

Prospects for further research are to applying similar approach of drift detection to more complex data structures, as well as a further study of SNN's ability to learn spatio-temporal patters in the data streams.

## REFERENCES

1. Ramírez-Gallego Sergio [et al.] A survey on Data Preprocessing for Data Stream Mining: Current status and future directions [Electronic resource], *Neurocomputing*, 2017, Vol. 239. Mode of access: https://doi.org/10.1016/j.neucom.2017.01.078. Title from screen.
2. Gerstner Wulfram [et al.] Spiking Neuron Models: Single Neurons, Populations, Plasticity [Electronic resource] [S. l. : s. n.], 2002. Mode of access: https://doi.org/10.1017/CBO9780511815706. Title from screen.
3. Davidson S., Furber S. Comparison of Artificial and Spiking Neural Networks on Digital Hardware [Electronic resource], *Frontiers in Neuroscience*, 2021, Vol. 15, P. 651141. Mode of access: https://doi.org/10.3389/fnins.2021.651141. Title from screen.
4. Kasabov N. NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data [Electronic resource], *Neural Networks*, 2014, Vol. 52. Mode of access: https://doi.org/10.1016/j.neunet.2014.01.006. Title from screen.
5. Jesús López Lobo [et al.] Spiking Neural Networks and online learning: An overview and perspectives, 2020.
6. Jesús López Lobo [et al.] Drift Detection over Non-stationary Data Streams Using Evolving Spiking Neural Networks [Electronic resource], *Studies in Computational Intelligence.* [S. l.], 2018, pp. 82–94. Mode of access: https://doi.org/10.1007/978-3-319-99626-4_8. Title from screen.
7. Hochreiter S., Schmidhuber J. {Long Short-Term Memory} [Electronic resource], *Neural Computation*, 1997, Vol. 9, No. 8, pp. 1735–1780. Mode of access: https://doi.org/10.1162/neco.1997.9.8.1735. Title from screen.
8. Cho Kyunghyun [et al.] Learning phrase representations using RNN encoder-decoder for statistical machine translation, *arXiv preprint arXiv:1406.1078*, 2014.
9. Voelker A., Kajic I., Eliasmith C. Legendre Memory Units: Continuous-Time Representation in Recurrent Neural Networks, [S. l.], 2019.
10. Fang Wei [et al.] Incorporating learnable membrane time constant to enhance learning of spiking neural networks, *Proceedings of the IEEE/CVF international conference on computer vision*, [S. l.], 2021, pp. 2661–2671.
11. Samant Rucha [et al.] A Systematic Ensemble Approach for Concept Drift Detector Selection in Data Stream Classifiers [Electronic resource], *International Journal of Engineering Trends and Technology*, 2022, Vol. 70, pp. 119–130. Mode of access: https://doi.org/10.14445/22315381/IJETT-V70I9P212. Title from screen.


DOI 10.15588/1607-3274-2024-4-8

12. Priya S, Annie R. S P. Ensemble framework for concept drift detection and class imbalance in data streams [Electronic resource], *Multimedia Tools and Applications*, 2024. pp. 1–15. Mode of access: https://doi.org/10.1007/s11042-024-18349-y. Title from screen.
13. Wu Yafeng [et al.]. An Adaptive Ensemble Framework for Addressing Concept Drift in IoT Data Streams [Electronic resource], 2023. Mode of access: https://doi.org/10.36227/techrxiv.23304461. Title from screen.
14. Dakic Kosta [et al.]Spiking Neural Networks for Detecting Satellite-Based Internet-of-Things Signals, *arXiv preprint arXiv:2304.03896*, 2023.
15. Régis Antônio S. Albuquerque [et al.] A decision-based dynamic ensemble selection method for concept drift, *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. [S. l.], 2019, pp. 1132–1139.
16. Yoonsik Shim [et al.] Unsupervised Learning in an Ensemble of Spiking Neural Networks Mediated by ITDP [Electronic resource], *PLOS Computational Biology*, 2016, Vol. 12, P. e1005137. Mode of access: https://doi.org/10.1371/journal.pcbi.1005137. Title from screen.
17. Deng L. The mnist database of handwritten digit images for machine learning research, *IEEE Signal Processing Magazine*, 2012, Vol. 29, No. 6, pp. 141–142.
18. Neculae G., Rhodes O., Brown G. Ensembles of Spiking Neural Networks, 2021.
19. Orchard Garrick [et al.] Converting static image datasets to spiking neuromorphic datasets using saccades, *Frontiers in neuroscience*, 2015, Vol. 9, P. 159859.
20. Amir Arnon [et al.]A Low Power, Fully Event-Based Gesture Recognition System [Electronic resource], 2*017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 7388–7397. Mode of access: https://api.semanticscholar.org/CorpusID:19496174. Title from screen.
21. Fu Q., Dong H. An ensemble unsupervised spiking neural network for objective recognition [Electronic resource], *Neurocomputing*, 2021, Vol. 419, pp. 47–58. Mode of access: https://doi.org/10.1016/j.neucom.2020.07.109. Title from screen.
22. Krizhevsky A. Learning multiple layers of features from tiny images, [S. l. : s. n.], 2009.
23. Street W. N., Kim Y. S. A streaming ensemble algorithm (SEA) for large-scale classification [Electronic resource], *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA, 2001, pp. 377–382. Mode of access: https://doi.org/10.1145/502512.502568. Title from screen.
24. Paszke Adam [et al.] Automatic differentiation in PyTorch, NIPS-W, [S. l.], 2017.
25. Fang Wei [et al.]SpikingJelly: An open-source machine learning infrastructure platform for spike-based intelligence [Electronic resource], *Science Advances*, 2023, Vol. 9, No. 40, P. eadi1480. Mode of access: https://doi.org/10.1126/sciadv.adi1480. – Title from screen.
26. Angelopoulos Angelos [et al.] Impact of Classifiers to Drift Detection Method: A Comparison / // Proceedings of the 22nd Engineering Applications of Neural Networks Conference / ed. by L. Iliadis [et al.]. Cham, 2021, pp. 399–410.
27. Suryawanshi Shubhangi [et al.] Adaptive windowing based recurrent neural network for drift adaption in non-stationary environment [Electronic resource], *Journal of Ambient Intelligence and Humanized Computing*, 2022, Vol. 14, pp. 1–15. Mode of access: https://doi.org/10.1007/s12652-022-04116-0. Title from screen.
28. Tavanaei Amirhossein [et al.] Deep learning in spiking neural networks, *Neural networks*, 2019, Vol. 111, pp. 47–63.

Received 24.07.2024.
Accepted 17.10.2024.

УДК 004.8

## АНСАМБЛЬ ПРОСТИХ СПАЙКОВИХ НЕЙРОННИХ МЕРЕЖ В ЯКОСТІ ДЕТЕКТОРІВ ДРЕЙФУ КОНЦЕПЦІЇ

**Бодянський Є. В.** – д-р техн. наук, професор, професор кафедри штучного інтелекту Харківського національного університету радіоелектроніки, Харків, Україна.

**Савенков Д. В.** – аспірант кафедри штучного інтелекту Харківського національного університету радіоелектроніки, Харків, Україна.

### АНОТАЦІЯ

**Актуальність.** У цій статті запропоновано новий підхід до виявлення дрейфу концепцій з використанням ансамблю простих спайкових нейронних мереж. Такий підхід використовує подієву природу та вбудовану здатність нейронів вивчати просторово-часові патерни, а ансамбль забезпечує додаткову робастність та масштабованість. Це може допомогти вирішити актуальну проблему обмеженості часових та обчислювальних ресурсів у задачах онлайн машинного навчання, особливо в дуже суворих середовищах, таких як IoT, які також мають інші переваги від використання шипінг-обчислень.

**Мета роботи.** Метою роботи є створення ансамблю простих спайкових нейронних мереж для роботи в якості детектора концептуального дрейфу в задачах інтелектуального аналізу потоків даних в Інтернеті.

**Метод.** Запропонований підхід в першу чергу базується на накопичувальній природі спайкових нейронних мереж, особливо негерметичних нейронів інтеграції-та-пострілу, які можна розглядати як одиниці пам'яті із затворами, де мембранна постійна часу $\tau_m$ є константою балансу між запам'ятовуванням та забуванням інформації. Реалізовано алгоритм навчання, який використовує неглибоку двошарову SNN, що використовує ознаки та мітки даних як вхідний шар, а другий шар складається з одного нейрона. Активація цього нейрона означає, що відбувся різкий дрейф. Крім того, така модель використовується як базова модель в ансамблі для покращення робастності, точності та масштабованості.

**Результати.** Ансамбль неглибоких двошарових SNN було реалізовано та навчено для виявлення різкого дрейфу концепції в потоці даних SEA. Ансамблю вдалося значно підвищити точність порівняно з базовою моделлю та досягти конкурентних результатів із сучасними передовими моделями.

**Висновки.** Результати показали життєздатність запропонованого рішення, яке не тільки забезпечує дешеве і конкурентоспроможне рішення для середовищ з обмеженими ресурсами, але і відкриває двері для подальших досліджень здатності спайкових нейромереж вивчати просторово-часові патерни в потоках даних та інших областях.

**КЛЮЧОВІ СЛОВА:** машинне навчання, онлайн навчання, спайкові нейронні мережі, дрейф концепцій, детектор дрейфу, штучні нейронні мережі, інтелектуальний аналіз потоку даних, штучний інтелект, негерметичний нейрон інтеграції-та-пострілу.


DOI 10.15588/1607-3274-2024-4-8

**ЛІТЕРАТУРА**

1. A survey on Data Preprocessing for Data Stream Mining: Current status and future directions [Electronic resource] / Sergio Ramírez-Gallego [et al.] // Neurocomputing. – 2017. – Vol. 239. – Mode of access: https://doi.org/10.1016/j.neucom.2017.01.078. – Title from screen.
2. Spiking Neuron Models: Single Neurons, Populations, Plasticity [Electronic resource] / Wulfram Gerstner [et al.]. – [S. l. : s. n.], 2002. – Mode of access: https://doi.org/10.1017/CBO9780511815706. – Title from screen.
3. Davidson S. Comparison of Artificial and Spiking Neural Networks on Digital Hardware [Electronic resource] / Simon Davidson, Steve Furber // Frontiers in Neuroscience. – 2021. – Vol. 15. – P. 651141. – Mode of access: https://doi.org/10.3389/fnins.2021.651141. – Title from screen.
4. Kasabov N. NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data [Electronic resource] / Nikola Kasabov // Neural Networks. – 2014. – Vol. 52. – Mode of access: https://doi.org/10.1016/j.neunet.2014.01.006. – Title from screen.
5. Spiking Neural Networks and online learning: An overview and perspectives / Jesús López Lobo [et al.]. – 2020.
6. Drift Detection over Non-stationary Data Streams Using Evolving Spiking Neural Networks [Electronic resource] / Jesús López Lobo [et al.] // Studies in Computational Intelligence. – [S. l.], 2018. – P. 82–94. – Mode of access: https://doi.org/10.1007/978-3-319-99626-4_8. – Title from screen.
7. Hochreiter S. {Long Short-Term Memory} [Electronic resource] / Sepp Hochreiter, Jürgen Schmidhuber // Neural Computation. – 1997. – Vol. 9, no. 8. – P. 1735–1780. – Mode of access: https://doi.org/10.1162/neco.1997.9.8.1735. – Title from screen.
8. Learning phrase representations using RNN encoder-decoder for statistical machine translation / Kyunghyun Cho [et al.] // arXiv preprint arXiv:1406.1078. – 2014.
9. Voelker A. Legendre Memory Units: Continuous-Time Representation in Recurrent Neural Networks / Aaron Voelker, Ivana Kajic, Chris Eliasmith. – [S. l.], 2019.
10. Incorporating learnable membrane time constant to enhance learning of spiking neural networks / Wei Fang [et al.] // Proceedings of the IEEE/CVF international conference on computer vision. – [S. l.], 2021. – P. 2661–2671.
11. A Systematic Ensemble Approach for Concept Drift Detector Selection in Data Stream Classifiers [Electronic resource] / Rucha Samant [et al.] // International Journal of Engineering Trends and Technology. – 2022. – Vol. 70. – P. 119–130. – Mode of access: https://doi.org/10.14445/22315381/IJETT-V70I9P212. – Title from screen.
12. S P. Ensemble framework for concept drift detection and class imbalance in data streams [Electronic resource] / Priya S, Annie R // Multimedia Tools and Applications. – 2024. – P. 1–15. – Mode of access: https://doi.org/10.1007/s11042-024-18349-y. – Title from screen.
13. An Adaptive Ensemble Framework for Addressing Concept Drift in IoT Data Streams [Electronic resource] / Yafeng Wu [et al.]. – 2023. – Mode of access: https://doi.org/10.36227/techrxiv.23304461. – Title from screen.
14. Spiking Neural Networks for Detecting Satellite-Based Internet-of-Things Signals / Kosta Dakic [et al.] // arXiv preprint arXiv:2304.03896. – 2023.
15. A decision-based dynamic ensemble selection method for concept drift / Régis Antônio S. Albuquerque [et al.] // 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). – [S. l.], 2019. – P. 1132–1139.
16. Unsupervised Learning in an Ensemble of Spiking Neural Networks Mediated by ITDP [Electronic resource] / Yoon-sik Shim [et al.] // PLOS Computational Biology. – 2016. – Vol. 12. – P. e1005137. – Mode of access: https://doi.org/10.1371/journal.pcbi.1005137. – Title from screen.
17. Deng L. The mnist database of handwritten digit images for machine learning research / Li Deng // IEEE Signal Processing Magazine. – 2012. – Vol. 29, No. 6. – P. 141–142.
18. Neculae G. Ensembles of Spiking Neural Networks / Georgiana Neculae, Oliver Rhodes, Gavin Brown. – 2021.
19. Converting static image datasets to spiking neuromorphic datasets using saccades / Garrick Orchard [et al.] // Frontiers in neuroscience. – 2015. – Vol. 9. – P. 159859.
20. A Low Power, Fully Event-Based Gesture Recognition System [Electronic resource] / Arnon Amir [et al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P. 7388–7397. – Mode of access: https://api.semanticscholar.org/CorpusID:19496174. – Title from screen.
21. Fu Q. An ensemble unsupervised spiking neural network for objective recognition [Electronic resource] / Qiang Fu, Hongbin Dong // Neurocomputing. – 2021. – Vol. 419. – P. 47–58. – Mode of access: https://doi.org/10.1016/j.neucom.2020.07.109. – Title from screen.
22. Krizhevsky A. Learning multiple layers of features from tiny images / Alex Krizhevsky. – [S. l. : s. n.], 2009.
23. Street W. N. A streaming ensemble algorithm (SEA) for large-scale classification [Electronic resource] / W. Nick Street, YongSeog Kim // Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – New York, NY, USA, 2001. – P. 377–382. – Mode of access: https://doi.org/10.1145/502512.502568. – Title from screen.
24. Automatic differentiation in PyTorch / Adam Paszke [et al.] // NIPS-W. – [S. l.], 2017.
25. SpikingJelly: An open-source machine learning infrastructure platform for spike-based intelligence [Electronic resource] / Wei Fang [et al.] // Science Advances. – 2023. – Vol. 9, No. 40. – P. eadi1480. – Mode of access: https://doi.org/10.1126/sciadv.adi1480. – Title from screen.
26. Impact of Classifiers to Drift Detection Method: A Comparison / Angelos Angelopoulos [et al.] // Proceedings of the 22nd Engineering Applications of Neural Networks Conference / ed. by L. Iliadis [et al.]. – Cham, 2021. – P. 399–410.
27. Adaptive windowing based recurrent neural network for drift adaption in non-stationary environment [Electronic resource] / Shubhangi Suryawanshi [et al.] // Journal of Ambient Intelligence and Humanized Computing. – 2022. – Vol. 14. – P. 1–15. – Mode of access: https://doi.org/10.1007/s12652-022-04116-0. – Title from screen.
28. Deep learning in spiking neural networks / Amirhossein Tavanaei [et al.] // Neural networks. – 2019. – Vol. 111. – P. 47–63.


DOI 10.15588/1607-3274-2024-4-8

p-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2024. № 4
e-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2024. № 4

UDC 004.383.8.032.26

# HARDWARE IMPLEMENTATION OF AN ANALOG SPIKING NEURON WITH DIGITAL CONTROL OF INPUT SIGNALS WEIGHING

**Gnilenko A. B.** – PhD, Associate Professor, Associate Professor of the Department of Electronic Computing Machinery, Oles Honchar Dnipro National University, Dnipro, Ukraine.

## ABSTRACT

**Context.** Significant challenges facing hardware developers of artificial intelligence systems force them to look for new non-standard architectural solutions. One of the promising solutions is the transition from von Neumann's classic architecture to neuromorphic architecture, which at the hardware level tries to imitate the work of the neural network of the human brain. A neuromorphic processor built as hardware implementation of a spiking neural network consists of a large number of elementary electronic circuits that structurally and functionally correspond to neurons. Thus, the design of hardware implementation of a spiking neuron as the basic building element of a neuromorphic processor is of great scientific interest.

**Objective.** The goal of the work is to design an analog spiking neuron hardware implementation with digital control of input signals by binary synaptic weighting coefficients.

**Method.** Designing is performed at the logical/schematic and topological levels of the design flow using modern tools of electronic design automation. All proposed schematic and layout solutions are verified and simulated using computer aided design tools to prove their functionality.

**Results.** The schematic and layout solutions have been developed and investigated for the hardware implementation of the spiking analog neuron with digital control of input signals by binary synaptic weighting coefficients to be the basic building element of a spiking neural network of the neuromorphic processor.

**Conclusions.** The proposed hybrid design of the spiking neuron hardware implementation benefits by combining the simplicity of analog signal processing methods in the neuron with digital control of the state of the neuron using binary weighting coefficients. The simulation results confirm the functionality of the obtained schematic/layout solutions and demonstrate the possibility of implementing logical functions inherent in the perceptron. The prospects for further research may include the design of hardware implementation for a spiking neural network core based on the developed schematic and layout solutions for the spiking neuron.

**KEYWORDS:** neuromorphic processor, spiking neural network, neuron, synaptic coefficient, layout design.

## ABBREVIATIONS

AI is artificial intelligence;
ANN is an artificial neural network;
SNN is a spiking neural network;
VLSI is very large-scale integration;
IC is integrated circuit;
CAD is computer aided design;
EDA is electronic design automation;
DRC is design rule check;
ERC is electrical rule check;
LVS is layout versus schematic;
LIF is leaky integrate-and-fire;
STDP is spike timing dependent plasticity;
GPU is a graphics processing unit;
MOS is metal-oxide-semiconductor;
Si-IPD is silicon-based integrated passive devices;
DC is a direct current.

## NOMENCLATURE

$V_{DD}$ is a positive power supply voltage for integrated circuits;

$V_{GND}$ is a zero voltage reference point in the circuit used as a baseline for measuring the voltage of other circuit components;

$V_{\text{base.}}$ is a reference voltage for the excitatory or inhibitory circuits of the synaptic input;

$V_{\text{offset}}$ is a decremental or incremental voltage for the excitatory or inhibitory circuits of the synaptic input;

$w_i$ is a synaptic weighting coefficient of the $i$-th synaptic input of the neuron.

## INTRODUCTION

The growing interest in building AI systems based on neural networks prompts the search for new architectural solutions for hardware, which would be more adequate to the methods of solving problems in the neural network basis. It becomes clear that using arrays of graphics processors for the hardware of AI systems has significant limitations in terms of energy efficiency, autonomy, mobility, scalability, etc., and can only be considered as a temporary solution. Therefore, in recent years, more and more attention has been drawn to neuromorphic computer systems that are built and function similarly to the biological neural network of the human brain. A neuromorphic processor in such a system is a VLSI circuit organized as a spiking neural network of computing elements that functionally correspond to neurons. The design of neuromorphic processors, their basic elements and the hardware implementation of neuromorphic calculations is an actual problem of modern computer engineering.

**The object of study** is the process of signal processing control in the hardware implementation of the analog spiking neuron.

**The subject of study** is the methods for the design and verification of the hardware implementation of the analog spiking neuron for a neuromorphic processor neural network.

**The purpose of the work** is to develop and study schematic and layout solutions for the hardware implementation of the analog spiking neuron with the digital


DOI 10.15588/1607-3274-2024-4-9

control of input signals by binary synaptic weighting coefficients.

## 1 PROBLEM STATEMENT

The key problem of neuromorphic computing is the design and fabrication of neuromorphic chips based on the hardware implementation of spiking neurons and spiking neural networks as VLSI circuits [1, 2]. The spiking neural network of a neuromorphic processor can be built based on the use of analog or digital circuits [3–6]. Both approaches have some advantages and drawbacks. But the best option would rather be a hybrid circuit as a combination of analog and digital parts. The optimal choice for the hardware implementation of a neuromorphic processor is the LIF model of the spiking neuron. In [7, 8], a simple implementation of the synapse of the LIF model of an analog spiking neuron was proposed, which uses circuits of parallel-connected MOS transistors that control the current at the output of the synapse to weigh the input signals. However, no method has been proposed to control the weighing of input signals based on the values of the weighting coefficients. That is why the problem is to design the hardware implementation of a spiking analog neuron with digital control of input signals weighing based on the synaptic input model described in [7]. The hardware implementation of the hybrid spiking neuron must be designed at the schematic and layout level of the VLSI design flow using CAD tool Tanner EDA with appropriate verification procedures.

## 2 REVIEW OF THE LITERATURE

Growing demands of modern AI systems reveal problems and limitation their hardware implementation based on traditional architecture solutions. Using supercomputers and GPU arrays as the hardware platform of AI systems results in the problems of huge energy consumption, limited scalability, and the lack of autonomy. In recent years, it has become obvious that the most promising way to solve these problems would be the transition to neuromorphic systems. Such systems at the hardware level implement the concept of spiking neural networks for the organization of massive parallel processing of information.

Spiking neural networks belong to the third generation of neural networks. SNN is a type of artificial neural network that mainly relies on methods of transmitting and processing information in the form in which they exist in the biological nervous system. As in its biological original, in a SNN, information is transmitted through connections between neurons using short pulses – spikes, which are generated by neurons when they are activated. In the biological nervous system, a spike corresponds to a nerve impulse.

An important feature of the SNN is the asynchronous nature of the work of neurons, due to which there is no constant flow of data across all synaptic connections between neurons, as in a conventional artificial neural network. Each neuron in SNN works independently of the others, responding only to the arrival of a spike from one of the pre-synaptic neurons by changing the membrane potential. This event-driven and asynchronous nature of the work makes it possible to build an energy-efficient hardware implementation of a spiking neural network as a neuromorphic chip.

The method of processing spikes in an artificial neuron depends on the neuron model used in the spiking neural network. There are various models [9] of the neuron, many of which have been implemented in hardware, which can be divided into biologically plausible, biologically inspired, integrate-and-fire models, and derivatives of the original McCulloch-Pitts model. Biologically plausible models such as Hodgkin-Huxley model [10] try to describe the physical processes in the biological neuron as accurately as possible in terms of ion's transferring mechanism through the neuron cell membrane. Biologically inspired models are simplified variants of the Hodgkin-Huxley model, for example, Fitzhugh-Nagumo [11], Hindmarsh-Rose [12] models, or the popular Izhikevich spiking neuron model [13], which accurately model the behavior of the biological neuron rather than its physical activity. Although those models are more computation-oriented and have a simpler hardware implementation than for the Hodgkin-Huxley model, they are still too complex to practically build neuromorphic systems based on them.

Less biologically realistic and more computationally simple are a set of integrate-and-fire models of the spiking neuron. These models and their more advanced variant – Leaky Integrate-and-Fire model offer a balance between the accuracy of the description of the neuron's behavior and the simplicity of calculations [14].

The dynamics of the LIF model of a neuron is illustrated in Fig. 1 [15]. Spikes from pre-synaptic neurons are first modulated by the weights of synaptic inputs, then currents at synaptic inputs are integrated into the membrane potential of the neuron, which exponentially decreases over time due to leakage. If the membrane potential exceeds the threshold value, the neuron generates an action potential in the form of a post-spike, after which the membrane potential decreases to the level of the resting potential.

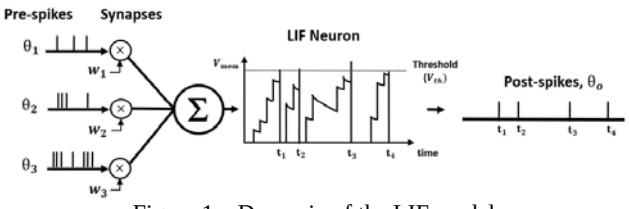



Figure 1 – Dynamic of the LIF model

The behavior of LIF model accurately imitates biological neural network dynamic. That is why this model of a spiking neuron is currently the most popular in the hardware implementation of neuromorphic systems.

Classical ANN training methods based on the back-propagation algorithm, are hardly applicable to spiking


DOI 10.15588/1607-3274-2024-4-9

neural networks, since the SNN is digital in nature, and in addition, backpropagation does not comply with the principle of biological plausibility. That is why there are two approaches to train SNN. First one is the method of Spike Timing Dependent Plasticity [16] based on the principle of locality [17]. According to this method the synapses which activate the post-synaptic neuron shortly after receiving spikes increase their weight, while the synapses that received spikes shortly after post-synaptic neuron activation decrease their weights. The STDP method is ready for the hardware implementation of learning mechanisms and makes it possible to train SNN directly on the neuromorphic chip. The second approach [18,19] applies the method of out-of-chip training that implies the construction of the classical ANN as a continuous analogue of the SNN and teaching this auxiliary network using backpropagation method with a host computer. After training, obtained synaptic weight coefficients are translated into synaptic memory on the neuromorphic chip.

For more than two decades of research in the field of neuromorphic systems and computing, there have been several significant projects, most of which continue to this day [17]. One of the first projects was SpiNNaker, launched in 2011 at the University of Manchester and continued as a part of the European Human Brain Project on the SpiNNaker 2 hardware platform [20]. The SpiNNaker is not a neuromorphic processor but it is rather a massively parallel computing system that was specially built for simulating a spiking neural networks to model human brain structures. The first industrially produced neuromorphic chip was the TrueNorth chip [21], created in 2014 by IBM under the auspices of the DARPA SyNAPSE program. The spiking neurons of the chip cores were built as simplified digital circuits that allow only addition and subtraction operations to be performed. Synaptic weight coefficients were coded with 2-bit values, that is why only out-of-chip training was possible using a different hardware platform with the translation of weight coefficient values obtained as training results into synaptic memory of the TrueNorth chip. The first neuromorphic chip with the ability to learn directly on the chip was the Liohi chip [22] created by Intel in 2018 using a 14 nm technological process, and the Liohi 2 chip that was presented in 2021 and was already made by a 7 nm technology. Synaptic weights in the Liohi chip can already be encoded with 8-bit values, which makes it possible to perform on-chip training. Synaptic weights are modified in the process of local training according to the rule of synaptic plasticity, which is formulated in the form of a simple formula with only addition and multiplication operations implemented in a set of microprograms. The Tianjic neuromorphic chip presented by Beijing Xinhua University in 2019 became the first hybrid chip that combines the architecture of the classical artificial neural network and the spiking neural network in one neuromorphic system. [23,24]. The Tianjic chip is not designed for on-chip training. Therefore, training is carried out on the platform of graphic processors, and the results of training are implemented in the synaptic memory of the chip.

There are other successful projects in the field of development of neuromorphic systems and their number is constantly increasing. The commercial success of artificial intelligence systems that we have observed in recent years and the apparent inability to support this success with adequate hardware platforms leads to the need for further research in the field of neuromorphic systems and their hardware implementation.

## 3 MATERIALS AND METHODS

Let's consider the block diagram of an analog spiking neuron shown in Fig. 2. The neuron circuit contains a Schmidt trigger that implements the threshold function of activation and a ring oscillator that generates spikes. A leaky integrator forms the membrane potential of the neuron, ensuring the accumulation of the potential to a threshold value upon the arrival of input signals and its gradual reduction to the resting potential in the absence of input spikes. The addition of input signals in the analog neuron is carried out as the usual sum of currents in an electric circuit.

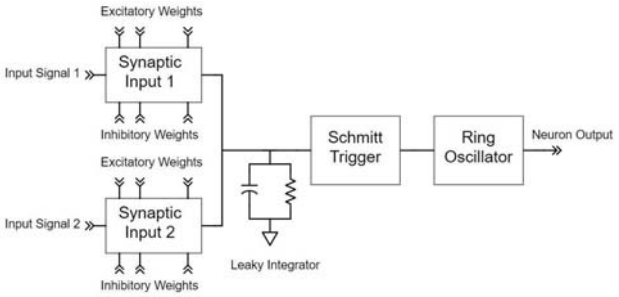



Figure 2 – Block diagram of an analog spiking neuron with two synaptic inputs [25]

But the most significant components of the neuron are the synaptic inputs, which produce weighted input signals that are amplified or attenuated according to the values of the synaptic weights. To weigh the input signal, the hardware implementation of the synapse has been proposed in [7,8] which uses excitation and inhibition circuits consisting of several MOS transistors connected in parallel, as shown in Fig. 3.

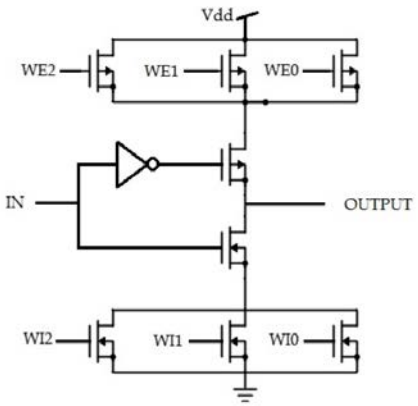



Figure 3 – Circuit diagram for the synaptic input [7]


DOI 10.15588/1607-3274-2024-4-9

By changing the voltage levels on the gates of the transistors in the excitation and/or inhibition circuits, the input signal can be weighted. This can be done by alternately switching these transistors to the subthreshold mode of operation, which allows us to gradually increase or decrease the current at the synapse output. Regardless of the number of transistors in the circuits, it is sufficient to use two voltage levels – $V_{base}$ and $V_{offset}$, where $V_{DD} \geq V_{base} > V_{offset}$ for the excitation circuit, and $V_{GND} \leq V_{base} < V_{offset}$ for the inhibition circuit, which are alternately applied to the gates and switch the transistors one by one in the subthreshold mode of operation. Three MOS transistors in the excitation and inhibition circuits, as shown in Fig. 3, provide a possibility to obtain four levels of current at the synapse output and, accordingly, four levels of the post-synaptic potential at the leaky integrator.

The selection of the post-synaptic potential level is carried out by switching the gates of the transistors in the excitation and/or inhibition circuits with one of two power sources – $V_{base}$ or $V_{offset}$. To control this selection for excitation and inhibition circuits consisted of three MOS transistors, 2-bit values can be used acting as synaptic weighting coefficients.

The circuit diagram for digital control of voltage on the gates of MOS transistors in the excitation and inhibition circuits is shown in Fig. 4, with voltage switches built on two transmission gates [26].

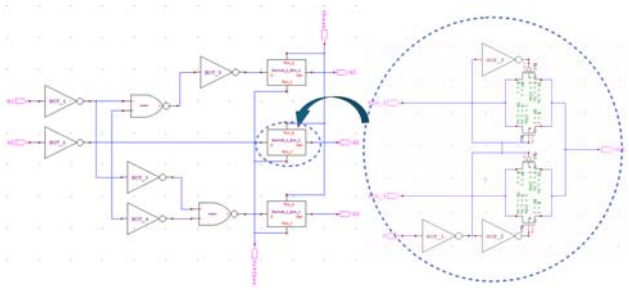


Figure 4 – Circuit diagram for the digital control of voltage

The layout design of the voltage control circuit is demonstrated in Fig. 5.

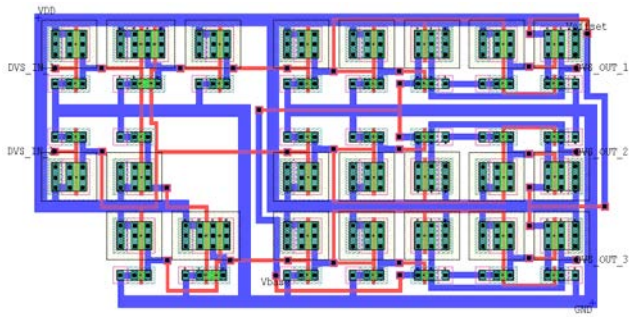



Figure 5 – Layout design of the voltage control circuit

The circuit diagram of the synaptic input with two blocks of digital control with excitatory and inhibitory weighting coefficients is shown in Fig. 6. Unlike a biological neuron, this hardware implementation allows us to use both excitatory and inhibitory weighting coefficients, which adds additional functional flexibility to such a neuron implementation.

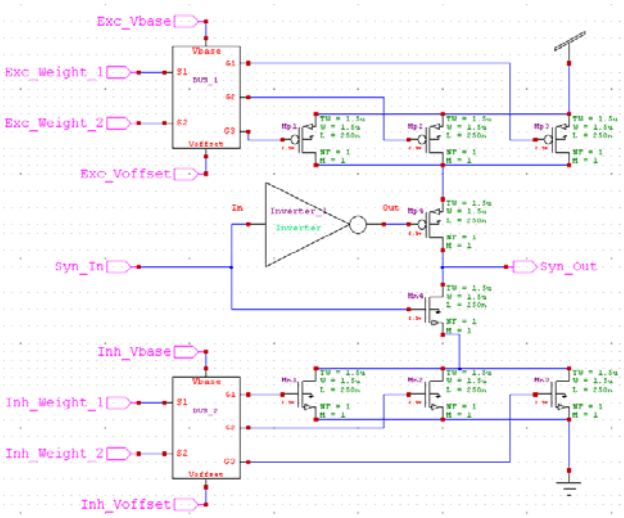



Figure 6 – Circuit diagram for the synaptic input with the digital control of input signal by synaptic weights

The advantage of the analog implementation of the neuron is the simplicity of adding new synaptic inputs to the neuron, since the summation of weighted input signals occurs naturally according to Kirchhoff's law as the sum of currents.

For the hardware implementation of the threshold activation function, a Schmitt trigger is used, based on the circuit on MOS transistors, as shown in Fig. 7. The Schmitt trigger converts the analog input signal of the membrane potential into a digital output signal that activates a spike generator if the membrane potential of the neuron exceeds the threshold level. A classic ring oscillator consisting of an odd number of inverters arranged in a ring is used as a spike generator.

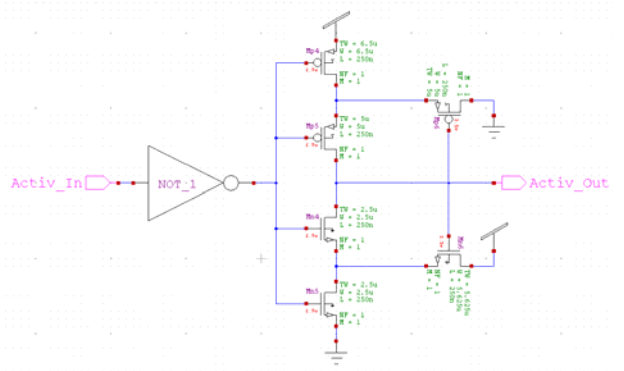



Figure 7 – Circuit diagram of the Schmitt trigger

The layout design of the neuron activator based on the Schmitt trigger and the generator of spikes based on the ring oscillator are shown in Fig. 8 and Fig. 9, respectively.

Based on the developed components of the spiking neuron, a neuron circuit with two synaptic inputs was designed in accordance with the block diagram shown in


DOI 10.15588/1607-3274-2024-4-9

Fig. 2. The layout design of the neuron with two synaptic inputs and the digital control of input signals weighing by binary synaptic weighting coefficients is shown in Fig. 10. It should be noted that capacitive and resistive elements of the leaky integrator are formed on the back side of the silicon wafer using 3D Silicon-based Integrated Passive Devices (Si-IPD) technology and through-silicon vias used to integrate passive components such as inductors, resistors and capacitors on a chip. The characteristics of those elements were directly added to the netlist exported from the neuron layout for its verification.

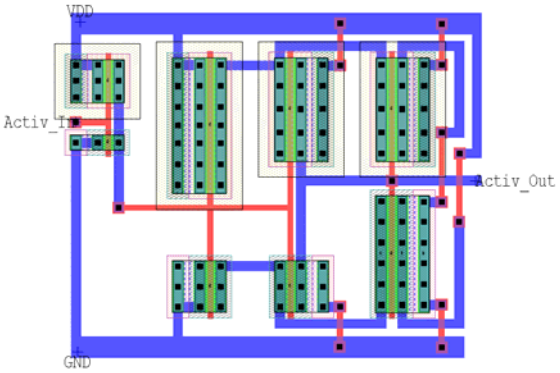



Figure 8 – Layout design of the Schmitt trigger

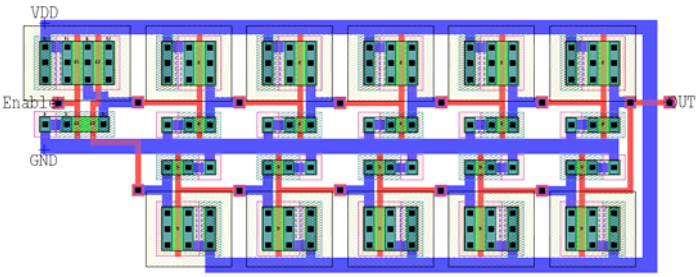



Figure 9 – Layout design of the ring oscillator

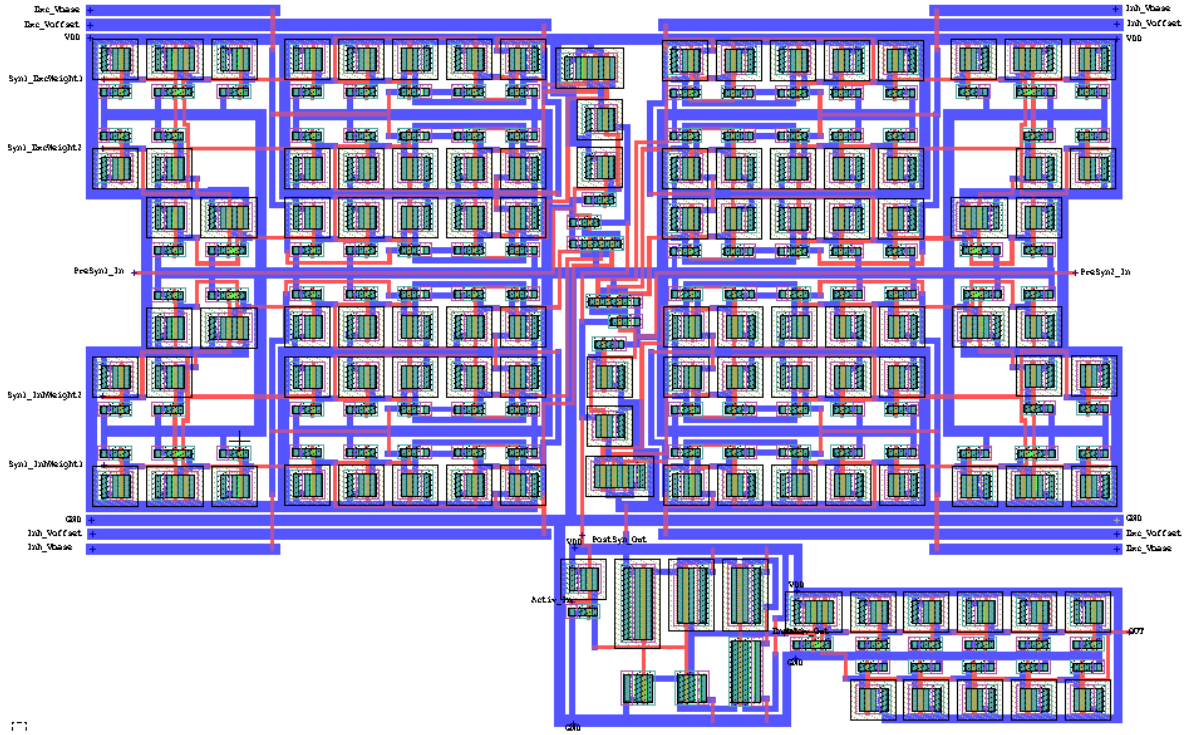


Figure 10 – Layout design of the analog spiking neuron with two synaptic inputs and digital control of input signals weighing

## 4 EXPERIMENTS

The spiking analog neuron hardware implementation is designed using Tanner EDA software which is a professional standard in the field of VLSI design. The Tanner EDA provides a complex software solution for the design and verification at the schematic and layout levels of the design flow for full-custom analog and mixed signal integrated circuits. The design flow can be demonstrated using the diagram shown in Fig. 11.

After defining circuit specifications, a logic/transistor circuit is designed at schematic level using S-Edit component of Tanner EDA. To simplify the design of larger structures, a symbolic designation (Symbol) is created and associated with the circuit. Simulation of the circuit at the schematic level is carried out to determine its functionality and, in case of errors, the circuit is revised. The verified circuit is exported to the netlist of schematic level.


DOI 10.15588/1607-3274-2024-4-9

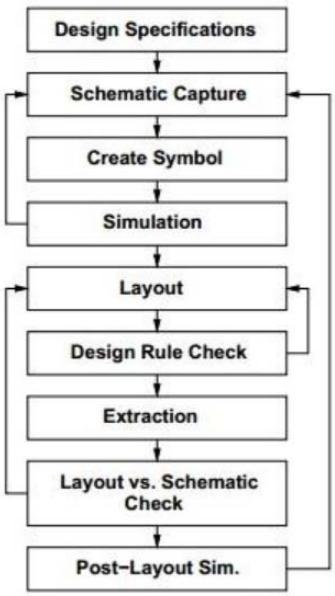



Figure 11 – Design flow diagram

According to the transistor circuit, a topological drawing (Layout) of the placement of diffusion areas of transistors, gates, interconnects, power buses, etc. on a silicon crystal is developed using L-Edit component of Tanner EDA. Layout verification is carried out for compliance with design rules and technological standards using Design Rule Check, (DRC) and Electrical Rule Check (ERC) procedures. If non-compliance with the design rules is detected, the layout design is corrected.

Next, the layout is exported to the netlist of the layout level and is verified by comparing the netlists of the schematic and layout levels using Layout versus Schematic Check (LVS) component. When inconsistencies are detected, the layout design is redone. At the final stage of design flow, Post Layout Simulation is carried out at the layout level to compare with that of schematic level. Simulation results at schematic and layout levels are obtained as waveform probing of signals at reference nodes of the circuit.

## 5 RESULTS

It is well known that a neuron with two inputs is an elementary perceptron capable of implementing simple logical functions. Let's investigate the functionality of proposed schematic solutions using the example of calculating logical functions.

We will apply periodic signals to the pre-synaptic inputs of the neuron, which provide all four binary combinations possible for a device with two inputs. Waveforms probing is registered at some reference nodes of the neuron as input signals, membrane potential of the neuron, activation signal and output signal of the neuron. Since synaptic weighting coefficients are two-bit binary values, it is easy to estimate the effect of all possible combinations of weighting coefficients of the synaptic inputs on the output signal of the neuron. Assume that the input signals are affected only by the excitatory weighting coefficients of the synaptic inputs. Then, for a neuron with two synaptic inputs, we have sixteen possible combinations of two-bit weighting coefficients $w_1$ and $w_2$.

Some simulation results for combinations of excitatory synaptic weighting coefficients which provide the implementation of different logical functions are shown in Fig. 12 as signal waveforms at reference nodes of the neuron.

The generalized result of modeling the neuron's processing of input signals for sixteen combinations of excitatory synaptic weighting coefficients with interpretation in terms of logical functions is shown in Table 1 where A and B stand for signals at the pre-synaptic inputs.

Table 1 – Logical functions generated at neuron's output

| $w_2$ | $w_1$ | Neuron Output |
|---|---|---|
| 00 | 00 | 0 |
| 00 | 01 | A AND B |
| 00 | 10 | A |
| 00 | 11 | A |
| 01 | 00 | A AND B |
| 01 | 01 | A AND B |
| 01 | 10 | A |
| 01 | 11 | A |
| 10 | 00 | B |
| 10 | 01 | B |
| 10 | 10 | A OR B |
| 10 | 11 | A OR B |
| 11 | 00 | B |
| 11 | 01 | B |
| 11 | 10 | A OR B |
| 11 | 11 | A OR B |

To implement the logical NOT function, it is necessary to use inhibitory weighting coefficients. The input signal to be inverted is applied to the first synaptic input of the neuron, while at the same time the second input is applied with DC signal whose level corresponds to $V_{DD}$ value. The signal at the first synaptic input is affected by inhibitory weighting coefficient only while the signal at the second synaptic input is neither amplified nor inhibited. The signal waveforms at the reference nodes of the neuron are shown in Fig. 13 for inhibitory weighting coefficient of the first synaptic input $w_1$=10 while the others weighting coefficients are zero.

Thus, the designed circuit of an analog neuron with digital control of input signals weighing using binary synaptic coefficients acts as an elementary perceptron and can implement logical functions AND, OR and NOT.


DOI 10.15588/1607-3274-2024-4-9

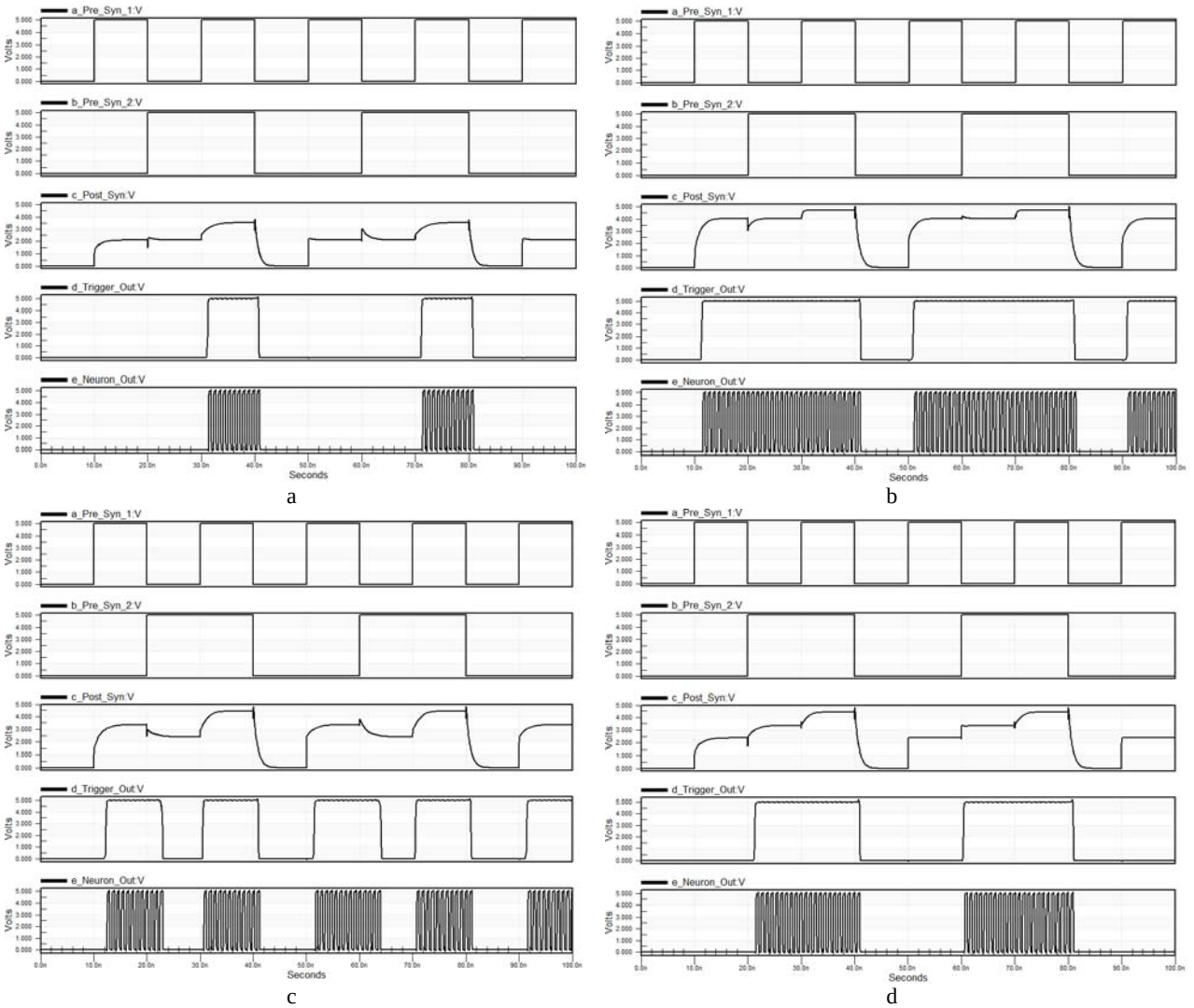



Figure 12 – Signal waveforms at reference nodes of the neuron for some combinations of weighting coefficients:
a – $w_1$=01, $w_2$=01, Out = A AND B; b – $w_1$=11, $w_2$=11, Out = A OR B; c – $w_1$=10, $w_2$=01, Out = A; d – $w_1$=01, $w_2$=10, Out = B

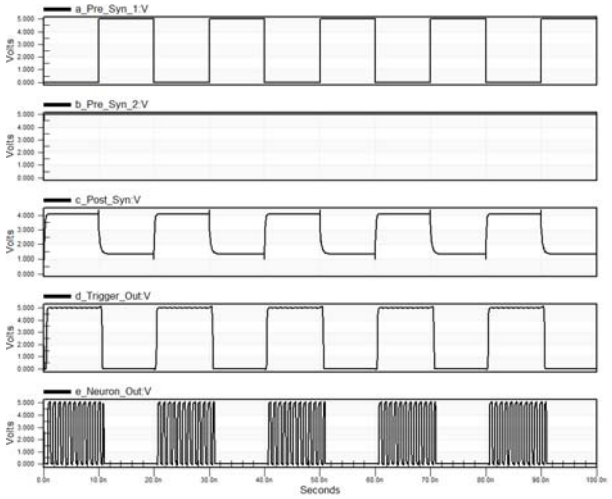



Figure 13 – Logical NOT function implementation using inhibitory weighting coefficient $w_1$=10

## 6 DISCUSSION

The spiking neuron is a node of a spiking neural network that forms a neuromorphic chip. That is why the spiking neuron can be considered as a basic building element for the neuromorphic processor. The main results of the paper are schematic and layout solutions for the hardware implementation of an analog spiking neuron with digital control of input signals weighing. Analog nature of the neuron allows us to simplify signal processing implementation, but the state of the neural network should be saved as digital values of weighting coefficients in distributed memory devices. This fact determines requirements to provide the digital control of input signals weighing using binary values of weighting coefficients.

The proposed solutions are based on the synaptic input schematic described in [7] which is the closest analogue to the considered hardware implementation. In [7], authors discus simple approach to weigh input signals using the excitation and inhibition circuits consisted of parallel connected *p*-type and *n*-type MOS transistors. But any methods to control the amplification and attenuation of


DOI 10.15588/1607-3274-2024-4-9

input signals using synaptic weighting coefficients have not been proposed. In the presented research for the synaptic input the schematic solution is developed to digitally control voltage on the gates of MOS transistors in the excitation and inhibition circuits. This allows us to control input signals of the neuron using binary values of the synaptic weighting coefficients.

The practical significance of the research is related to the layout design of the analog spiking neuron. This design can be used as the basic building block of a semiconductor intellectual property core to produce a neuromorphic chip as an application-specific integrated circuit.

The schematic solutions for all neuron blocks and the neuron as whole are verified by the simulation using CAD tools with obtaining waveforms of signals at reference nodes of circuits. The layout solutions are verified with DRC, ERC, LVS checks and post-layout simulations.

The simulation results obtained for all combinations of two-bit weighting coefficients demonstrate how the neuron processes input signals to calculate logical function which is the classical problem for the perceptron. The weighting coefficients for a specific problem can be found as the result of off-line learning and then translated to two-bit binaries.

The advantage of the proposed hybrid solution for the spiking neuron consists in the simplicity of neuron circuit scaling by adding new synaptic input blocks. The schematics of the neuron core and synaptic input blocks remain the same with adding new inputs to the neuron. That allows us to easily build neuron circuits with multiple inputs for a spiking neuron network of the neuromorphic processor.

The simplicity of scaling the neuron schematic and layout solutions opens up prospects for further research in the direction of developing spiking neural network structures of a neuromorphic processor.

## CONCLUSIONS

The problem of controlling the weighing of input signals for an analog spiking neuron using binary synaptic weighting coefficients has been solved.

**The scientific novelty** of obtained results is that the digital method of controlling the weighing of input signals for an analog spiking neuron is proposed. The hybrid design of the neuron improves the previously developed schematic of the synaptic input making possible to control input signals of the analog neuron using binary values of weighting coefficients. This allows us to use analog nature of signal processing by neurons while the state of neurons is determined by binary values of weighting coefficients stored in memory devices.

**The practical significance** of obtained results is that the layout is designed for the analog spiking neuron with digital controlling the weighing of input signals by binary synaptic weighting coefficients. This layout design can be used as the SIP core basic building block to produce a spiking neural network of a neuromorphic chip.

**Prospects for further research** are to develop and study structures of SNN for a neuromorphic processor based on the designed spiking analog neuron with the digitally controlled weighing of input signals.

## REFERENCES

1. Indiveri G., Linares-Barranco B., Hamilton T. J. et al. Neuromorphic silicon neuron circuits, *Frontiers in neuroscience*, 2011, Vol. 5, Art. 73, pp. 1–23. DOI: 10.3389/fnins.2011.00073
2. Sun J. CMOS and memristor technologies for neuromorphic computing applications, *Technical Report No. UCB/EECS-2015-219*, 2015, P. 31. [Electronic Resource]. Access mode: https://www2.eecs.berkeley.edu/ Pubs/ TechRpts/2015/EECS-2015-219.pdf
3. Joubert A., Belhad B., Temam O. et al. Hardware spiking neurons design: analog or digital? *International Joint Conference on Neural Networks: The 2012 International Conference, Brisbane, 10–15 June 2012: proceedings.* IEEE, 2012, pp. 1–5. DOI:10.1109/IJCNN.2012.6252600
4. Yammenavar B. D., Gurunaik V. R., Bevinagidad R. N. et al. Design and analog VLSI implementation of artificial neural network, *International Journal of Artificial Intelligence & Applications*, 2011, Vol. 2, No. 3, pp. 96–109. DOI:10.5121/ijaia.2011.2309
5. Forssell M. Hardware Implementation of Artificial Neural Networks. [Electronic Resource]. Access mode: https://users.ece.cmu.edu/~pgrover/teaching/files/ NeuromorphicComputing.pdf
6. Shinde J. R., Salankar S. VLSI implementation of neural network, *Current Trends in Technology and Science*, 2015, Vol. 4, No. 3, pp. 515–524
7. Yellamraju S., Kumari S., Girolkar S. et al. Design of various logic gates in neural networks, *IEEE India Conference: 2013 Annual IEEE India Conference, Mumbai, 13–15 December 2013: proceedings*. IEEE, 2013, pp. 1–5. DOI:10.1109/INDCON.2013.6725879
8. Liu B., Konduri S., Minnich R. et al. Implementation of pulsed neural networks in CMOS VLSI technology, *Proceedings of the 4th WSEAS International Conference on Signal Processing, Robotics and Automation, 13 February*, 2005, Art. No. 20, pp. 1–8
9. Schuman C. D., Potok T. E., Patton R. M. et al. A survey of neuromorphic computing and neural networks in hardware, arXiv:1705.06963v1 [cs.NE], 2017, pp. 1–88
10. Hodgkin A. L., Huxley A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 1952, Vol. 117, No. 4, pp. 500–544. DOI: 10.1113/jphysiol.1952.sp004764
11. Binczak S., Jacquir S., Bilbault J.-M. et al. Experimental study of electrical FitzHugh-Nagumo neurons with modified excitability, *Neural Networks*, 2006, Vol. 19, No. 5, pp. 684–693. DOI:10.1016/j.neunet.2005.07.011
12. Hayati M., Nouri M., Abbott D. et al. Digital multiplierless realization of two-coupled biological hindmarsh–rose neuron model, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2016, Vol. 63, No. 5, pp. 463–467. DOI:10.1109/TCSII.2015.2505258
13. Izhikevich E. M. Simple model of spiking neurons, *IEEE Transactions on neural networks*, 2003, Vol. 14, No. 6, pp. 1569–1572. DOI: 10.1109/ TNN.2003.820440
14. Lu S., Xu F. Linear leaky-integrate-and-fire neuron model based spiking neural networks and its mapping relationship to deep neural networks, *Frontiers in neuroscience*, 2022, Vol. 16, 16:857513. DOI: 10.3389/ fnins.2022.857513
15. Lee C., Sarwar S., Panda P. et al. Enabling spike-based backpropagation for training deep neural network architec-


DOI 10.15588/1607-3274-2024-4-9

tures, *Frontiers in neuroscience*, 2020, Vol. 14, 14:119. DOI: 10.3389/fnins.2020.00119
16. Diehl P. U., Cook M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity, *Frontiers in Computational Neuroscience*, 2015, Vol. 9. DOI: 10.3389/fncom.2015.00099
17. Ivanov D., Chezhegov A., Kiselev M.et al. Neuromorphic artificial intelligence systems, *Frontiers in Neuroscience*, 2022, Vol. 16, 16:959626. DOI: 10.3389/ fnins.2022.959626
18. Bohte S. M., Kok J. N., Poutre H. L. et al. Error-backpropagation in temporally encoded networks of spiking neurons, *Neurocomputing*, 2001, No. 48, pp. 17–37. DOI:10.1016/S0925-2312(01)00658-0
19. Esser S. K., Appuswamy R., Merolla P. A. et al. Backpropagation for Energy-Efficient Neuromorphic Computing, *Proc. of. Neural Information Processing Systems: 28 Int. Conference*. Montreal, 7 December 2015, Vol. 1, pp. 1117–1125
20. Höppner S., Yan Y., Dixius A. et al. The SpiNNaker 2 processing element architecture for hybrid digital neuromorphic computing, *arXiv preprint arXiv:2103.08392*, 2021. DOI: 10.48550/arXiv.2103. 08392
21. Merolla P. A., Arthur J. V., Alvarez-Icaza R.et al. A million spiking-neuron integrated circuit with a scalable communication network and interface, *Science*, 2014, Vol. 345, No. 6197, pp. 668–673. DOI: 10.1126/science. 1254642
22. Davies M., Srinivasa N., Lin T.-H.et al. Loihi: a neuromorphic manycore processor with on-chip learning, *IEEE Micro*, 2018, Vol. 38, No. 1, pp. 82–99. DOI: 10.1109/MM.2018.112130359
23. Pei J., Deng L., Song S. et al. Towards artificial general intelligence with hybrid tianjic chip architecture, *Nature*, 2019. Vol. 572, pp. 106–111. DOI: 10.1038/s41586-019-1424-8
24. Wang G., Ma S., Wu Y. et al. End-to-end implementation of various hybrid neural networks on a cross-paradigm neuromorphic chip, *Frontiers in neuroscience*, 2021, Vol. 15, 15:615279. DOI: 10.3389/fnins.2021.615279.r
25. Gnilenko A.B. Hardware implementation design of a spiking neuron, *System technologies*, 2021, Vol. 132, No. 1, pp. 116–123. DOI: 10.34185/1562-9945-1-132-2021-10
26. Gnilenko A.B. Layout design of a synaptic input with digitally controlled weight coefficients for a hardware implementation of an artificial spiking neuron, *System technologies*, 2023, Vol. 144, No. 1, pp. 76–82. DOI: 10.34185/1562-9945-1-144-2023-10

Received 19.08.2024.
Accepted 16.10.2024.

УДК 004.383.8.032.26

## АПАРАТНА РЕАЛІЗАЦІЯ АНАЛОГОВОГО ІМПУЛЬСНОГО НЕЙРОНА З ЦИФРОВИМ КЕРУВАННЯМ ЗВАЖУВАННЯМ ВХІДНИХ СИГНАЛІВ

**Гниленко О. Б.** – канд. фіз.-мат. наук, доцент, доцент кафедри електронних обчислювальних машин Дніпровського національного університету імені Олеся Гончара, Дніпро, Україна.

### АНОТАЦІЯ

**Актуальність.** Значні виклики, що постають перед розробниками апаратного забезпечення систем штучного інтелекту, змушують шукати для реалізації таких систем нові нестандартні архітектурні рішення. Одним із таких перспективних рішень є перехід від класичної архітектури фон Неймана до нейроморфної архітектури, яка на апаратному рівні намагається імітувати роботу нейронної мережі людського мозку. Нейроморфний процесор, побудований як апаратна реалізація імпульсної нейронної мережі, складається з великої кількості елементарних електронних схем, які структурно та функціонально відповідають нейронам. Тому, проєктування апаратної реалізації імпульсного нейрона як основного будівельного елементу нейроморфного процесора представляє собою значний науковий та практичний інтерес.

**Мета роботи.** Метою роботи є розробка апаратної реалізації аналогового імпульсного нейрона з цифровим керуванням зважуванням вхідних сигналів двійковими синаптичними ваговими коефіцієнтами.

**Метод.** Проєктування виконується на схемотехнічному та топологічному рівнях наскрізного маршруту проєктування інтегральних схем з використанням сучасних засобів автоматизації проєктування електронних пристроїв. Для підтвердження функціональності усіх запропонованих схемотехнічних та топологічних рішень проведено їх верифікацію та моделювання засобами автоматизованого проєктування,.

**Результати.** Розроблено та досліджено схемотехнічні та топологічні рішення для апаратної реалізації аналогового імпульсного нейрона з цифровим керуванням зважуванням вхідних сигналів двійковими синаптичними ваговими коефіцієнтами як основного елемента побудови імпульсної нейронної мережі нейроморфного процесора.

**Висновки.** Запропонована гібридна конструкція апаратної реалізації імпульсного нейрона має переваги завдяки поєднанню простоти аналогових методів обробки сигналів в нейроні з цифровим керуванням станом нейрона за допомогою двійкових вагових коефіцієнтів. Результати моделювання підтверджують функціональність отриманих схемотехнічних та топологічних рішень і демонструють можливість реалізації логічних функцій, притаманних персептрону. Перспективи подальших досліджень можуть включати розробку апаратної реалізації ядра імпульсної нейронної мережі нейроморфного процесора на основі розроблених схемотехнічних та топологічних рішень для імпульсного нейрона..

**КЛЮЧОВІ СЛОВА:** нейроморфний процесор, імпульсна нейронна мережа, нейрон, синаптичний коефіцієнт, топологічне проєктування.


DOI 10.15588/1607-3274-2024-4-9

**ЛІТЕРАТУРА**

1. Neuromorphic silicon neuron circuits / [G. Indiveri, B. Linares-Barranco, T. J. Hamilton et al.] // Frontiers in neuroscience. – 2011. – Vol. 5, Art. 73. – P. 1–23. DOI: 10.3389/fnins.2011.00073
2. Sun J. CMOS and memristor technologies for neuromorphic computing applications / J. Sun // Technical Report No. UCB/EECS-2015-219. – 2015.– P. 31. – [Electronic Resource]. – Access mode: https://www2.eecs.berkeley.edu/Pubs/ TechRpts/2015/EECS-2015-219.pdf
3. Hardware spiking neurons design: analog or digital? / [A. Joubert, B. Belhad, O. Temam et al.] // International Joint Conference on Neural Networks: The 2012 International Conference, Brisbane, 10–15 June 2012: proceedings. – IEEE, 2012. – P. 1–5. DOI:10.1109/IJCNN.2012.6252600
4. Design and analog VLSI implementation of artificial neural network / [B. D. Yammenavar, V. R. Gurunaik, R. N. Bevinagidad et al.] // International Journal of Artificial Intelligence & Applications. – 2011. – Vol. 2, No. 3. – P. 96–109. DOI:10.5121/ijaia.2011.2309
5. Forssell M. Hardware Implementation of Artificial Neural Networks / M. Forssell // [Electronic Resource]. – Access mode: https://users.ece.cmu.edu/~pgrover/teaching/files/NeuromorphicComputing.pdf
6. Shinde J. R. VLSI implementation of neural network / J. R. Shinde, S. Salankar // Current Trends in Technology and Science. – 2015. – Vol. 4, No. 3. – P. 515–524.
7. Design of various logic gates in neural networks / [S. Yellamraju, S. Kumari, S. Girolkar et al.] // IEEE India Conference: 2013 Annual IEEE India Conference, Mumbai, 13–15 December 2013: proceedings. – IEEE, 2013. – P. 1–5. DOI:10.1109/INDCON.2013.6725879
8. Implementation of pulsed neural networks in CMOS VLSI technology / [B. Liu, S. Konduri, R. Minnich et al.] // Proceedings of the 4th WSEAS International Conference on Signal Processing, Robotics and Automation, 13 February 2005. – Art. No. 20. – P. 1–8
9. A survey of neuromorphic computing and neural networks in hardware / [C. D. Schuman, T. E. Potok, R. M. Patton et al.] // arXiv:1705.06963v1 [cs.NE]. – 2017.– P. 1–88
10. Hodgkin A. L. A quantitative description of membrane current and its application to conduction and excitation in nerve / A. L. Hodgkin, A. F. Huxley // The Journal of physiology. – 1952. – Vol. 117, No. 4. – P. 500–544. DOI: 10.1113/jphysiol.1952.sp004764
11. Experimental study of electrical FitzHugh-Nagumo neurons with modified excitability / [S. Binczak, S. Jacquir, J.-M. Bilbault et al.] // Neural Networks. – 2006. – Vol. 19, No. 5. – P. 684–693. DOI:10.1016/j.neunet.2005.07.011
12. Digital multiplierless realization of two-coupled biological hindmarsh–rose neuron model / [M. Hayati, M. Nouri, D. Abbott et al.] // IEEE Transactions on Circuits and Systems II: Express Briefs. – 2016. – Vol. 63, No. 5. – P. 463–467. DOI:10.1109/TCSII.2015.2505258
13. Izhikevich E. M. Simple model of spiking neurons / E. M. Izhikevich // IEEE Transactions on neural networks. – 2003. – Vol. 14, No. 6. – P. 1569–1572. DOI: 10.1109/TNN.2003.820440
14. Lu S. Linear leaky-integrate-and-fire neuron model based spiking neural networks and its mapping relationship to deep neural networks / S. Lu, F. Xu // Frontiers in neuroscience. – 2022. – Vol. 16, 16:857513. DOI: 10.3389/fnins.2022.857513
15. Enabling spike-based backpropagation for training deep neural network architectures / [C. Lee, S. Sarwar, P. Panda et al.] // Frontiers in neuroscience. – 2020. – Vol. 14, 14:119. DOI: 10.3389/fnins.2020.00119
16. Diehl P. U. Unsupervised learning of digit recognition using spike-timing-dependent plasticity / P. U. Diehl, M. Cook // Frontiers in Computational Neuroscience. – 2015. – Vol. 9. DOI: 10.3389/fncom.2015.00099
17. Neuromorphic artificial intelligence systems / [D. Ivanov, A. Chezhegov, M. Kiselev et al.] // Frontiers in Neuroscience. – 2022. – Vol. 16, 16:959626. DOI: 10.3389/fnins.2022.959626
18. Error-backpropagation in temporally encoded networks of spiking neurons / [S. M. Bohte, J. N. Kok, H. L. Poutre et al.] // Neurocomputing. – 2001. – No. 48. – P. 17–37. DOI:10.1016/S0925-2312(01)00658-0
19. Back-propagation for Energy-Efficient Neuromorphic Computing / [S. K. Esser, R. Appuswamy, P. A. Merolla et al.] // Proc. of. Neural Information Processing Systems*:* 28 Int. Conference, Montreal, 7 December 2015.– Vol. 1. – P. 1117–1125.
20. The SpiNNaker 2 processing element architecture for hybrid digital neuromorphic computing / [S. Höppner, Y. Yan, A. Dixius et al.] // arXiv preprint arXiv:2103.08392. – 2021. DOI: 10.48550/arXiv.2103. 08392
21. A million spiking-neuron integrated circuit with a scalable communication network and interface / [P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza et al.] // Science. – 2014. – Vol. 345, No. 6197. – P. 668–673. DOI: 10.1126/science. 1254642
22. Loihi: a neuromorphic manycore processor with on-chip learning / [M. Davies, N. Srinivasa, T.-H. Lin et al.] // IEEE Micro. – 2018. – Vol. 38, No. 1. – P. 82–99. DOI: 10.1109/MM.2018.112130359
23. Towards artificial general intelligence with hybrid tianjic chip architecture / [J. Pei, L. Deng, S. Song et al.] // Nature. – 2019. – Vol. 572. – P.106–111. DOI: 10.1038/s41586-019-1424-8
24. End-to-end implementation of various hybrid neural networks on a cross-paradigm neuromorphic chip / [G. Wang, S. Ma, Y. Wu et al.] // Frontiers in neuroscience. – 2021. – Vol. 15, 15:615279. DOI: 10.3389/fnins.2021.615279.r
25. Gnilenko A. B. Hardware implementation design of a spiking neuron / A. B. Gnilenko // System technologies. – 2021. – Vol. 132, No. 1. – P. 116–123. DOI: 10.34185/1562-9945-1-132-2021-10
26. Gnilenko A. B. Layout design of a synaptic input with digitally controlled weight coefficients for a hardware implementation of an artificial spiking neuron / A. B. Gnilenko // System technologies. – 2023. – Vol. 144, No. 1. – P. 76–82. DOI: 10.34185/1562-9945-1-144-2023-10


DOI 10.15588/1607-3274-2024-4-9

UDC 004.8

# IMPACT OF PREPROCESSING AND COMPARISON OF NEURAL NETWORK ENSEMBLE METHODS FOR SEGMENTATION OF THE THORACIC SPINE IN X-RAY IMAGES

**Koniukhov V. D.** – Postgraduate student, A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine, Kharkiv, Ukraine.

**Morgun O. M.** – PhD, Director of the"Laboratory of X-ray Medical Equipment" LTD, Kharkiv, Ukraine.

**Nemchenko K. E.** – Dr. Sc., Head of the Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

## ABSTRACT

**Context.** Automatic segmentation of medical images plays an important role in the process of automating the detection of various diseases in the spine and the use of radiography is the most accessible means of predicting diseases. Over the years many studies have been conducted on the topic of image segmentation. One of the many methods for improving image segmentation is the use of neural network ensembles.

**Objective.** The aims of this study were to investigate the impact of preprocessing and compare the main methods of neural network ensembles and their effect on the segmentation of the thoracic region, in this study the area was considered which consists of the vertebrae: Th8, Th9, Th10, Th11.

**Method.** To begin with, the influence of preprocessing of X-ray images was considered, which included the following methods: histogram equalization for contrast enhancement, contrast-limited adaptive histogram equalization, logarithmic transform method, median filter, Gaussian filter, and bilateral filter. To study the influence of neural network ensemble on segmentation quality, several methods were used. Averaging method – a simple half-averaging method. Weighted averaging method – an improved version of the averaging method which uses weights for each network, the higher the network weight, the greater its influence on averaging. Method of cumulative averaging – a modified averaging method in which each ensemble receives an averaged image, after which all the results of the ensembles are averaged. Bagging – method of averaging networks trained on different data, *n* networks are used, the training sample is divided into *n* parts, and each neural network is trained on its own subset of data, as a result, the averaging method is used for predictions. Averaging method for a large number of networks – in this method, 100 neural networks were trained, after which the averaging method was used. Method of averaging mask shapes – this method uses a distance transform to average multiple masks into one shape average.

**Results.** It was investigated that the use of different methods of image preprocessing does not guarantee an improvement in the quality of segmentation of the spine region on X-ray images, but even on the contrary worsens the quality of segmentation. Different methods of combining predictions of neural network ensembles were considered, which made it possible to find out the pros and cons of specific methods for the task of segmentation of X-ray images.

**Conclusions.** The experiments conducted allowed us to conclude that the use of any preprocessing methods should not be used for segmentation of X-ray images. Also, due to a large number of architectures and methods for combining predictions, the behavior of ensemble methods was studied, which will allow us to further determine the necessary approach for segmentation of X-ray images. Further study of the weighted averaging method and the mask shape averaging method will make it possible to improve the obtained result and achieve even greater success in segmentation.

**KEYWORDS:** machine learning; image recognition; neural network; image segmentation, computer vision.

## ABBREVIATIONS

BF is a bilateral filter;

CLAHE is a Contrast-Limited Adaptive Histogram Equalization;

DICE is a Dice-Sørensen coefficient;

GF is a Gaussian filter;

HE is a histogram equalization;

IoU is an Intersection over Union;

LT is a logarithmic transformation;

MF is a median filter.

## NOMENCLATURE

$A$ is a total number of networks in an ensemble;

$acc(S,P)$ is a method for determining the similarity of two images;

$B_M$ is a block width;

$B_N$ is a block height;

$c$ is a constant that scales the value after a logarithmic transformation;

$dt()$ is a function that calculates for each pixel the distance to the nearest zero pixel;

$FN$ is a false negative;

$FP$ is a false positive;

$g(I)$ is a method of combining predictions from different networks;

$I$ is an image matrix;

$I_{i,j}$ is a matrix element at position $(i, j)$;

$k$ is a half size of filter window;

$M$ is a width;

$N$ is a height;

$n_a$ is the number of architectures in an ensemble;

$n_{one}$ is a number of networks of the same architecture;

$n_e$ is a number of ensembles;

$n_n$ is a number of networks in an ensemble;


DOI 10.15588/1607-3274-2024-4-10

$P$ is a sequence of ground truth masks;
$p_i$ is a prediction of $i$-th model;
$S$ is a sequence of predictions of different networks;
$TP$ is true positive;
$v$ is a value of pixel;
$v_m$ is a value of one metric out of three;
$X$ is a pixel sequence;
$x$ is an input image;
σ is a standard deviation that determines the degree of blurring;
$\sigma_s$ is a degree of spatial smoothing;
$\sigma_r$ is an intensity smoothing degree;
$\sim mask$ is an inverted mask.

## INTRODUCTION

Automation of the spine segmentation process can significantly improve automatic diagnostics of diseases that require precise vertebrae selection. In the absence of a radiologist or his workload, a doctor who needs a radiologist's opinion can make a conclusion himself using the obtained results of automatic segmentation.

A large number of diseases of the spine require a better study of ways to improve its segmentation. The causes of Andersen's lesion are not completely clear, one of the theories accepted today is that primary inflammations are part of ankylosing spondylitis [1]. Cryptococcosis is an infectious disease, the cells of development of which can be vertebrae [2]. Fractures at the level of the 3rd and 4th vertebrae can injure the esophagus. As a result of ankylosing spondylitis, there is also a possibility of damage to the esophagus [3].

By using neural network ensembles, it is possible to improve the accuracy of X-ray image segmentation. Since the quality of X-ray images depends on many factors, this complicates the segmentation process. Analysis and comparison of different neural network ensemble methods can help determine the statistical pattern and select the best algorithm.

**The object of study** is the process of constructing ensembles of neural networks for segmenting the vertebral region. Different prediction-averaging algorithms are used to construct an ensemble of neural networks. The use of different averaging methods gives different results, which affects the result of using ensembles. The use of ensembles can cause an ambiguous effect, which in turn leads to the need to study these methods for segmenting the lumbar region since different methods have different effects on different objects.

**The subject of study** is methods of averaging ensembles of neural networks and preprocessing methods. Existing ensemble methods are considered in the context of certain tasks. The use of averaging algorithms provides different results for different objects. The study of existing methods and the development of new ones for segmentation of the vertebral region is a necessary part of the field of studying the segmentation of a certain region of the spine to compare their behavior with methods in another area.

**The purpose of the work** is to study the effect of preprocessing of vertebral images and compare neural network ensemble methods to determine the best one, the use of which will provide a guaranteed improvement in the segmentation of medical images.

## 1 PROBLEM STATEMENT

For segmentation of X-ray images, images of different quality are used. The images used can be represented as $I$, a two-dimensional matrix with the size $M \times N$. Each pixel of the image $I_{i,j}$ has a value from 0 to 255.

An ensemble of neural networks consists of $n$ networks. The creation of ensembles implies the use of neural networks of varying accuracy, which in turn, with a large number of neural networks, worsens the accuracy of segmentation. To solve this problem, it is necessary to supplement existing methods for obtaining ensemble results with new methods or modify existing ones. There are various methods for combining the results of neural network predictions, the most famous of which is the pixel averaging method. For each method that combines ensemble predictions, images are required:

$$S = \{I_i\}_{i=1}^{n}.$$

Due to the problems faced by various ensemble methods, it is necessary to define a method for which the following condition will be satisfied:

$$acc(g_i(S), P) > acc(g_j(S), P),$$

where $i, j \in \{1,2...,n\}$ and $i \neq j$.

## 2 REVIEW OF THE LITERATURE

Vertebral segmentation plays an integral role in disease diagnosis, preoperative preparation, and subsequent observation. Different quality of X-ray images requires the development of different approaches to extracting the necessary features from images. One of the reasons for obtaining poor-quality images is artefacts in the form of improper exposure. One study has shown that improper breathing technique significantly affects the quality of X-ray images [4]. In different countries, radiologists have different attitudes towards poor-quality images and images with poor image criteria. Non-compliance with image criteria, as well as poor image quality, are reasons for rejecting this image [5]. The dilemma of the impact of radiation on the patient and improving image quality forces doctors to make different decisions. After all, an increase in the dose entails an improvement in image quality, but at the same time, the risk for the patient increases. This issue creates the problem of finding a compromise to select the required radiation dose [6].

The use of machine learning methods for medical image segmentation is a very popular and in-demand approach nowadays. The use of this technique has signifi-


DOI 10.15588/1607-3274-2024-4-10

cantly helped to improve the quality of identifying various types of objects in X-ray images. In the process of improving neural networks and their modernization, various approaches have been proposed, one of which is an ensemble of neural networks. Thanks to ensembles of neural networks, it has become possible to combine less accurate classifiers, resulting in more accurate classifiers [7]. Deep learning models have one main problem – this is the need for a large amount of data and setting up optimal hyperparameters to achieve minimal error. The article [8] discusses the use of several ensemble methods for further application in a wide range of areas. Tests conducted in the article [9] showed that the resulting decrease in Boosting performance was due to overtraining in the presence of noise, which negatively affects the averaging result.

When considering ensembles of neural networks as an approach to improving the quality of segmentation, we should not forget about the methods that are based on the use of only one network. Studies of which have proven their ability to improve the search for specific objects in images. For example, a two-stage method using positioning of lumbar vertebrae and their subsequent segmentation due to the use of several networks at different stages: U-Net and XUnet, showed good performance [10]. The choice between using one network or several is not always obvious. In the article [11], the authors use only one network, since they claim that this approach tries to minimize redundancy in order to reduce the complexity of the network and reduce the training time. They note that this result can be obtained without sacrificing accuracy. The use of a two-branch multi-scale attention module that extracts the necessary information needed for segmentation of the vertebrae and the selection of key information in feature maps was proposed in the article [12]. In the paper [13], the authors proposed a vertebral segmentation method that segments one row of vertebrae as one individual spine object without training data using only manual identification for at least one vertebra. In the second step, they merge the shape prior to the segmentation flow of individual vertebrae. A method for localizing the lumbar spine using YOLOv5 and then passing the localized vertebrae through HED-U-Net to obtain the vertebrae and their edges was proposed in the paper [14].

Most of the studies on medical image segmentation use either neural networks specially designed for a specific task or general-purpose medical networks such as U-Net. The use of U-Net and deep learning for segmentation of the lumbar spine in MRI images demonstrated high segmentation accuracy [15]. Combining spinal canal segmentation using deep learning and morphological operations to solve the redundancy problems and improve the segmentation accuracy was proposed in the paper [16]. As can be seen from the extensive use of single neural networks for vertebrae segmentation, which shows good results, the use of ensembles of neural networks can further improve the results already obtained.

## 3 MATERIALS AND METHODS

Any machine learning method starts with data preparation. In some studies, on image segmentation you can find a stage of initial image preparation which may include: histogram equalization, median filter, Gaussian filter, etc. To begin with, it is worth considering these methods, because, in the case of medical X-ray images, they may not provide a positive result, but on the contrary, provide a negative result. Since the use of the same Gaussian filter can lead to deterioration in the quality of the boundaries.

The first method worth considering is the well-known histogram equalization algorithm. First, we need to build an initial histogram that will count the number of pixels for each intensity:

$$H(i) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} = \begin{cases} 1, I(x,y) = i \\ 0, otherwise \end{cases}.$$

The next step is to normalize the histogram:

$$H_{norm}(i) = \frac{H(i)}{M \times N}.$$

The third stage is the definition of the cumulative function:

$$CDF(i) = \sum_{j=0}^{i} H_{norm}(j).$$

Function to align each pixel:

$$I_{new}(x,y) = 255 \times CDF(I(x,y)).$$

Finally, we need to align all the pixels:

$$I_{new} = \{I_{new}(x,y) | 0 \le x < M, 0 \le y < N\}.$$

The second method for improving the contrast of CLAHE. The adaptive method differs from the usual one in that it calculates several histograms at once, each of the histograms corresponds to a separate section of the input image.

First of all, the core of the block is determined $B_M \times B_N$. Then for each block *b(i,j)* need to create a histogram $H_{i,j}(v)$. To prevent excessive contrast enhancement for each histogram, each block is limited by a threshold *T*. This is done in order to redistribute all values above a given threshold among the remaining values: $H'_{i,j}(v) = \min(H_{i,j}(v), T)$.

The next step is to determine the cumulative function:

$$CDF_{i,j}(v) = \sum_{n=0}^{v} H'_{i,j}(n).$$


DOI 10.15588/1607-3274-2024-4-10

And at the very end, we need to update the pixel values:

$$I_{new}(x,y) = round\left(\frac{CDF_{i,j}(I(x,y)) - \min(CDF_{i,j})}{(B_m \times B_n) - \min(CDF_{i,j})} \times 255\right).$$

The third method is the logarithmic transformation method. This method is used to highlight details with a low contrast level. The method looks like this:

$$I_{new}(x,y) = c \cdot \log(1 + I(x,y)).$$

The fourth method is the median filter. This method is designed to remove unwanted noise by dividing the image into windows in which all pixel values are grouped, after which the median value is determined. First, we need to define a window $W$ with the size $M \times N$, after which it is necessary to extract the values of all pixels from this window:

$$X = \{I(x_i, y_j) | 1 \le i < M, 1 \le j < N\}.$$

After which we will get a sorted array in ascending order $X_{sorted}$. Let's extract the median value $Med = median(X_{sorted})$, after which we write down the new value $I_{new}(x,y) = Med$.

The fifth method is Gaussian filter, which allows us to remove noise from an image by blurring it. First, the kernel is determined:

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

The value of the new pixel after applying this filter can be written as:

$$I_{new}(x,y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} G(i,j) \cdot I(x+i, y+j).$$

Bilateral filter is used in image processing to remove noise and smooth the image. The filter is calculated using the following formula:

$$I_{new}(x,y) = \frac{1}{W(x,y)} \sum_{i=-k}^{k} \sum_{j=-k}^{k} I(x+i, y+j) \cdot G_s(i,j) \cdot G_r(I(x+i, y+j) - I(x,y)).$$

Gaussian distribution:

$$G_s(i,j) = \exp\left(-\frac{i^2 + j^2}{2\sigma_r^2}\right).$$

Gaussian distribution is also used:

$$G_r(\Delta I) = \exp\left(-\frac{(\Delta I)^2}{2\sigma_r^2}\right).$$

The normalization coefficient has the following form:

$$W(x,y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} G_s(i,j) \cdot G_r(I(x+i, y+j) - I(x,y)).$$

After it became possible to obtain good results using image segmentation with machine learning, a new method was proposed – an ensemble of neural networks. By combining the predictions of several networks, it became possible to improve the performance of the model. One of the most famous methods that is used to combine the predictions of neural networks is the averaging method. This method has the following form:

$$Mean(n_n) = \frac{1}{n_n} \sum_{i=1}^{n_n} p_i.$$

The next ensemble method that was used was the weighted averaging method. The main difference from the averaging method is that each model is assigned a weight – this is a certain coefficient that is multiplied by the predictions in order to strengthen or weaken the final result of the network. This coefficient can be selected by different features, for example: DICE, IoU, and Recall. The formula for weighted averaging:

$$Weighted = \frac{\sum_{i=1}^{n_n} w_i p_i}{\sum_{i=1}^{n_n} w_i}.$$

If using the averaging method can give a better result, then we can also assume that if we create an averaging method for already averaged results, we can improve the indicators even more. The idea of this method is that we need $n$ results of averaging different ensembles, after which we average these results using the following formula:

$$CumulativeMean = \frac{1}{n_e} \sum_{i=1}^{n_e} Mean_i(n_n).$$

The last prediction fusion method worth considering is the proposed mask shape averaging method. It is based on the idea of transforming the distance of $n$ masks. For each mask, the distance from each pixel of the binary image to the nearest zero pixel is calculated. For zero pixels, the distance will be zero. The following formula displays the distance calculation for one mask:

$$d = dt(mask) - dt(\sim mask).$$

General formula for calculating the distance for all masks:

$$d_{all} = \sum_{i=1}^{n} d(mask_i).$$


DOI 10.15588/1607-3274-2024-4-10

In order to obtain the resulting mask, we sum up the obtained distances and if the sum of the values is greater than 0, then the pixel acquires the values 1 and 0 otherwise:

$$mask_{res} = d_{all} > 0.$$

## 4 EXPERIMENTS

For this study, open-source images were used [17]. The dataset consisted of 1098 chest X-ray images in lateral projection. The X-ray images depicted men and women. 1020 images were selected for training, the remaining 78 images were used for testing. There were 183 original images, and the remaining 915 were the result of augmentation. The following augmentation operations were performed: 1) random rotation in degrees [–15, 15]; 2) random shift in percentage vertically and horizontally [–10, 10]; 3) random scaling in percent [0.8, 1.2]; 4) random change in brightness [0.8, 1.2]; 5) random change of contrast [0.75, 1.5]. All images were reduced to the same resolution of 512×512 pixels. Figure 1 demonstrates an example of the image used and its masks that were in the dataset.

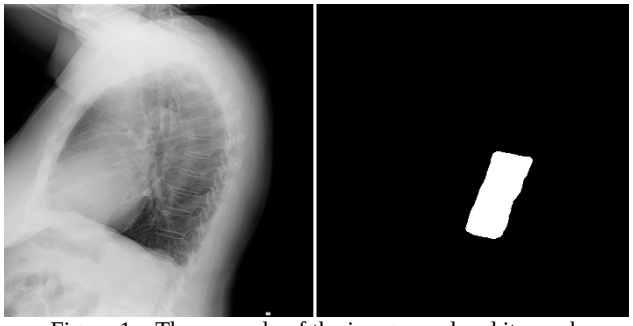


Figure 1 – The example of the image used and its mask

To study the diverse behavior of ensembles, it was decided to train ten models to study their behavior. The list of models used: Fcn8Mobilenet, Fcn8Resnet50, Fcn8Vgg, MobilenetUnet, Pspnet50, Pspnet101, Resnet50Pspnet, Resnet50Segnet, Resnet50Unet, VggUnet. Each used neural network architecture was trained five times. The number of epochs was set to 150, the batch size was eight, and EarlyStopping was used to prevent overfitting, which stopped training if the validation loss value did not improve over ten epochs.

To compare the predictions of neural networks and true masks, the Dice-Sørensen coefficient was used. Which serves as a binary measure of similarity and is expressed by the formula:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}.$$

Task 1. To study the meaning of using preprocessing methods, five Fcn8Mobilenet neural networks were used. Different filters were applied to each test image: histogram equalization, CLAHE, logarithmic transformation, median filter, Gaussian filter, and bilateral filter. After one of the above filters was applied to the image, a prediction was made for it from five networks.

Task 2. First, the averaging method was investigated. The main goal was to determine in which cases the best result can be achieved, due to the combination of which networks the model performance can be improved. First, ensembles of one architecture with different numbers of networks in the ensemble were used. Then, the possibility of combining different architectures with different numbers of networks in ensembles was considered. In the first case, for ensembles of one architecture, the following number of networks in one ensemble was used: 2, 3, 4, 5. In the second case of a combination of different architectures, the number of networks of one architecture in the ensemble was also: 2, 3, 4, 5, but the total number of networks in the ensemble was calculated as follows: $A=n_a \times n_{one}$. It is assumed that combining different architectures in an ensemble can have a good effect on the segmentation process due to different feature extraction and their combination. Each architecture will be assigned a code name to make it easier to write down combinations of names in the table. Architecture code names: A – Fcn8Mobilenet, B – Fcn8Resnet50, C – Fcn8Vgg, D – MobilenetUnet, E – Pspnet50, F – Pspnet101, G – Resnet50Pspnet, H – Resnet50Segnet, I – Resnet50Unet, J – VggUnet. Due to the large number of variations of network combinations, all results cannot be displayed in the table, so only a part will be displayed.

Task 3. The weighted averaging method requires a careful selection of weights for each network. In this case, it was proposed to use the following metrics: DICE, Precision, and Recall. Precision indicates the proportion of positive predictions that were specified correctly from all positive cases. The Precision formula is as follows:

$$Precision = \frac{TP}{TP + FP}.$$

The third metric, Recall, shows how many true positive cases were correctly predicted:

$$Recall = \frac{TP}{TP + FN}.$$

The weight for each network was obtained as follows:

$$w = \frac{v_m^{10}}{\sum_{i=1}^{n} v_m^{10}}.$$

Task 4. To study the cumulative averaging method, ten architectures were used, for each architecture four ensembles were used. The result of this algorithm was the combination of the results of the four ensembles. For each cumulative ensemble, averaging ensembles with the same number of networks were used, namely: an ensemble of


DOI 10.15588/1607-3274-2024-4-10

two networks, an ensemble of three networks, an ensemble of four networks and an ensemble of five networks.

Task 5. To implement this method, the training sample was divided into five parts. After that, five Fcn8Mobilenet networks were trained. Each network was trained on different data, so all networks had different images during the training process, which should have expanded the capabilities of the model for image segmentation. After the networks were trained, they were tested in four ensembles.

Task 6. In all previous studies in this article, a maximum of five networks of the same architecture were used, in this method it is proposed to check the influence of a large number of up to 100 networks of the same architecture on the segmentation result. Fcn8Mobilenet was trained 100 times on the same data.

Task 7. The last method that was used to combine the predicted masks was the method of averaging the shapes of masks based on the transformed distance. For its study, ten architectures were used and ensembles of 2, 3, 4, and 5 networks were created. Visually, the work of the algorithm can be seen in Figure 2. Where the white mask is the true mask of the image, the multi-colored contours are the contours of the masks used for averaging, the gold contour is the contour of the averaged mask.

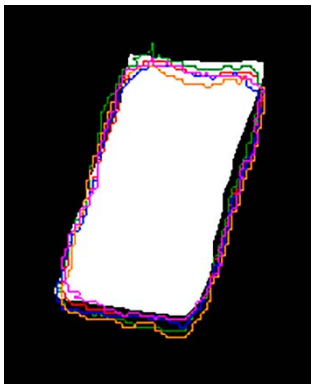


Figure 2 – The example of the method of averaging mask shapes with distance transformation

## 5 RESULTS

All trained neural networks are shown in Table 1. For greater clarity when comparing with ensemble methods, maximum, minimum and average values are also added.

The obtained results of the study of the application of image preprocessing methods are shown in Table 2. Only for two cases out of 30 a positive result was obtained, in the remaining cases these methods could not positively affect the image, which in turn did not lead to any improvement in the quality of segmentation, but on the contrary, worsened it. And even in these two cases, the result improved insignificantly. The application of these methods can have a good result in other areas, but in the study of X-ray images, they only harm. These methods blur or make the boundaries less clear, as a result of which the extraction of the necessary features suffers.

It was assumed that with the increase in the number of networks in the ensemble, DICE should increase, but this did not happen. Moreover, in most cases, with the increase in the number of networks, the DICE value decreased. Table 3 shows that in 12 cases out of 40 using averaging, it was possible to achieve a positive result compared to single networks. Only in the case of Resnet50Pspnet was it possible to achieve a positive result for all four ensembles. Fcn8Mobilenet was next with three positive results, and Pspnet50 closes the list with two successful ensembles, all other positive architectures had only one ensemble.

Using different architectures to build ensembles gave a positive result in the case of using nine and ten architectures simultaneously, which is demonstrated in Table 4. Combining a smaller number of networks in an ensemble had different results, and to achieve good results, it was necessary to use a large number of ensembles. It can be said with confidence that the simultaneous use of networks with a DICE difference greater than 0.05 will not bring a good result. Otherwise, a positive result could be seen with both a small and large difference in this value.

Table 5 shows the results of weighted averaging, in the process of studying which 90 cases were identified, using this method 22 positive results were obtained compared to the best single networks of the corresponding architectures. The best result was demonstrated by the Resnet50Pspnet architecture, which was able to achieve seven positive results out of nine, which is an excellent result.

The study of the use of the cumulative averaging method revealed that this method can improve the similarity of predictions and true masks compared to the best results of single networks in five cases out of ten. As for the comparison of this method and the averaging method, this method was able to surpass the competitive method in two cases out of ten. The results can be found in Table 6.

Training five networks of the same architecture on different data sets allowed us to minimize errors due to the ability to respond differently to input data. Using different data sets makes it possible to reduce the correlation of model errors. Table 7 shows the DICE for five trained networks. Studying the data provided in Table 8, we can conclude that in three out of four cases, using such an ensemble had a better result compared to the maximum result of a single network.

The use of a large number of networks of the same architecture in the ensemble is shown in Table 9. The expected increase in the result at each step of increasing the number of networks in the ensemble did not occur, but nevertheless, the best result was achieved using 100 networks. Although it is worth considering that the weight of 100 trained networks was 252 GB.

The last method considered was able to show the best result of all. 40 ensemble cases were considered, in 20 cases this method was better than maximal single networks and in 29 cases this method outperformed the averaging method, this can be seen in Table 10.


DOI 10.15588/1607-3274-2024-4-10

Table 1 – Similarity measure for trained neural networks

| Architecture name | Network number | | | | | Statistics | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Min | Max | Avg |
| Fcn8Mobilenet | 0.9179 | 0.9142 | 0.9061 | 0.9035 | 0.8933 | 0.8933 | 0.9179 | 0.9070 |
| Fcn8Resnet50 | 0.9539 | 0.9397 | 0.9272 | 0.9218 | 0.9188 | 0.9188 | 0.9539 | 0.9323 |
| Fcn8Vgg | 0.9457 | 0.9135 | 0.8896 | 0.8885 | 0.88442 | 0.8844 | 0.9457 | 0.9043 |
| MobilenetUnet | 0.9119 | 0.9078 | 0.9045 | 0.8780 | 0.8755 | 0.8755 | 0.9119 | 0.8956 |
| Pspnet50 | 0.9338 | 0.9121 | 0.9091 | 0.9040 | 0.8897 | 0.8897 | 0.9338 | 0.9097 |
| Pspnet101 | 0.9480 | 0.9358 | 0.8978 | 0.8767 | 0.8446 | 0.8446 | 0.9480 | 0.9006 |
| Resnet50Pspnet | 0.8938 | 0.8933 | 0.8920 | 0.8865 | 0.8553 | 0.8553 | 0.8938 | 0.8842 |
| Resnet50Segnet | 0.9420 | 0.9154 | 0.8941 | 0.8934 | 0.8788 | 0.8788 | 0.9420 | 0.9048 |
| Resnet50Unet | 0.9315 | 0.9157 | 0.9120 | 0.9115 | 0.9043 | 0.9043 | 0.9315 | 0.9150 |
| VggUnet | 0.9151 | 0.9067 | 0.8955 | 0.8849 | 0.8757 | 0.8757 | 0.9151 | 0.8956 |

Table 2 – Similarity measure for image preprocessing results

| Method name | Network number | | | | | Statistics | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Min | Max |
| HE | 0.8998 | 0.8971 | 0.8831 | 0.8871 | 0.8804 | 0.8804 | 0.8998 |
| CLAHE | 0.9206 | 0.8872 | 0.9037 | 0.8949 | 0.8772 | 0.8772 | 0.9206 |
| LT | 0.9108 | 0.9041 | 0.8771 | 0.9094 | 0.8595 | 0.8595 | 0.9108 |
| MF | 0.8997 | 0.8928 | 0.8890 | 0.8894 | 0.8860 | 0.8860 | 0.8997 |
| GF | 0.9156 | 0.9023 | 0.9026 | 0.9003 | 0.8933 | 0.8933 | 0.9156 |
| BF | 0.8730 | 0.8795 | 0.8684 | 0.8676 | 0.86062 | 0.8606 | 0.8795 |

Table 3 – Ensembles of neural networks of the same architecture

| Architecture name | Number of networks in the ensemble | | | | Statistics | |
|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | Min | Max |
| Fcn8Mobilenet | 0.9194 | 0.9192 | 0.9136 | 0.9185 | 0.9136 | 0.9194 |
| Fcn8Resnet50 | 0.9449 | 0.9493 | 0.9449 | 0.9466 | 0.9449 | 0.9493 |
| Fcn8Vgg | 0.9312 | 0.9372 | 0.9260 | 0.9280 | 0.9268 | 0.9372 |
| MobilenetUnet | 0.9104 | 0.9144 | 0.9057 | 0.9047 | 0.9047 | 0.9144 |
| Pspnet50 | 0.9208 | 0.9407 | 0.9360 | 0.9324 | 0.9208 | 0.9407 |
| Pspnet101 | 0.9382 | 0.9513 | 0.9330 | 0.9474 | 0.9330 | 0.9513 |
| Resnet50Pspnet | 0.8943 | 0.9090 | 0.9062 | 0.9049 | 0.8943 | 0.9090 |
| Resnet50Segnet | 0.9261 | 0.9352 | 0.9295 | 0.9289 | 0.9261 | 0.9352 |
| Resnet50Unet | 0.9261 | 0.9352 | 0.9295 | 0.9289 | 0.9261 | 0.9352 |
| VggUnet | 0.9062 | 0.9124 | 0.9146 | 0.9147 | 0.9062 | 0.9147 |

Table 4 – Ensembles of neural networks of different architectures

| Combinations | Number of networks of same architecture in ensemble | | | | | Statistics | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Min | Max |
| A+B | 0.9337 | 0.9383 | 0.9440 | 0.9399 | 0.9411 | 0.9383 | 0.9440 |
| A+B+C | 0.9524 | 0.9494 | 0.9467 | 0.9443 | 0.9432 | 0.9432 | 0.9524 |
| A+B+C+D | 0.9480 | 0.9468 | 0.9449 | 0.9423 | 0.9399 | 0.9399 | 0.9480 |
| A+B+C+D+E | 0.9496 | 0.9485 | 0.9470 | 0.9447 | 0.9428 | 0.9428 | 0.9496 |
| A+B+C+D+E+F | 0.9510 | 0.9510 | 0.9490 | 0.9473 | 0.9458 | 0.9458 | 0.9510 |
| A+B+C+D+E+F+G | 0.9525 | 0.9496 | 0.9484 | 0.9451 | 0.9437 | 0.9437 | 0.9525 |
| A+B+C+D+E+F+G+H | 0.9519 | 0.9509 | 0.9478 | 0.9466 | 0.9449 | 0.9449 | 0.9519 |
| A+B+C+D+E+F+G+H+I | 0.9549 | 0.9513 | 0.9490 | 0.9472 | 0.9463 | 0.9463 | 0.9549 |
| A+B+C+D+E+F+G+H+I+J | 0.9541 | 0.9512 | 0.9489 | 0.9474 | 0.9464 | 0.9464 | 0.9541 |
| A+D | 0.9138 | 0.9216 | 0.9269 | 0.9247 | 0.9210 | 0.9138 | 0.9269 |
| A+E | 0.9278 | 0.9321 | 0.9331 | 0.9348 | 0.9387 | 0.9278 | 0.9387 |
| A+G | 0.9042 | 0.9164 | 0.9222 | 0.9237 | 0.9240 | 0.9042 | 0.9240 |
| A+J | 0.9315 | 0.9323 | 0.9361 | 0.9311 | 0.9312 | 0.9311 | 0.9361 |
| A+K | 0.9177 | 0.9253 | 0.9248 | 0.9216 | 0.9240 | 0.9177 | 0.9253 |
| C+E | 0.9471 | 0.9450 | 0.9465 | 0.9441 | 0.9335 | 0.9335 | 0.9471 |
| C+F | 0.9529 | 0.9493 | 0.9447 | 0.9429 | 0.9431 | 0.9429 | 0.9529 |
| C+H | 0.9468 | 0.9467 | 0.9427 | 0.9433 | 0.9416 | 0.9416 | 0.9468 |
| C+J | 0.9519 | 0.9429 | 0.9458 | 0.9421 | 0.9378 | 0.9378 | 0.9519 |
| D+E+J | 0.9445 | 0.9402 | 0.9401 | 0.9413 | 0.9395 | 0.9395 | 0.9445 |
| D+E+K | 0.9368 | 0.9397 | 0.9388 | 0.9360 | 0.9354 | 0.9354 | 0.9397 |
| D+E+G | 0.9388 | 0.9369 | 0.9373 | 0.9329 | 0.9316 | 0.9316 | 0.9388 |
| D+E+G+J | 0.9407 | 0.9428 | 0.9434 | 0.9413 | 0.9393 | 0.9393 | 0.9434 |
| D+E+G+K | 0.9373 | 0.9388 | 0.9373 | 0.9358 | 0.9345 | 0.9345 | 0.9388 |
| D+K+E+J | 0.9393 | 0.9425 | 0.9423 | 0.9418 | 0.9417 | 0.9393 | 0.9425 |


DOI 10.15588/1607-3274-2024-4-10

Table 5 – Weighted averaging

| Architecture name | Results | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Precision | | | Recall | | | DSC | | |
| | $N_n$ | | | | | | | | |
| | 3 | 4 | 5 | 3 | 4 | 5 | 3 | 4 | 5 |
| Fcn8Mobilenet | 0.9141 | 0.9181 | 0.9185 | 0.9141 | 0.9168 | 0.9169 | 0.9141 | 0.9181 | 0.9185 |
| Fcn8Resnet50 | 0.9308 | 0.9465 | 0.9466 | 0.9308 | 0.9465 | 0.9466 | 0.9308 | 0.9465 | 0.9466 |
| Fcn8Vgg | 0.9312 | 0.9403 | 0.9385 | 0.9132 | 0.9333 | 0.9280 | 0.9312 | 0.9403 | 0.9385 |
| MobilenetUnet | 0.9144 | 0.9144 | 0.9144 | 0.9031 | 0.9118 | 0.9047 | 0.9031 | 0.9118 | 0.9047 |
| Pspnet50 | 0.8981 | 0.9372 | 0.9324 | 0.8981 | 0.9378 | 0.9329 | 0.8980 | 0.9372 | 0.9324 |
| Pspnet101 | 0.9382 | 0.9382 | 0.8726 | 0.9520 | 0.9513 | 0.8723 | 0.9382 | 0.9504 | 0.8722 |
| Resnet50Pspnet | 0.8860 | 0.9074 | 0.9049 | 0.8943 | 0.9082 | 0.9050 | 0.8860 | 0.9074 | 0.9049 |
| Resnet50Segnet | 0.8983 | 0.8939 | 0.9095 | 0.9321 | 0.9321 | 0.9321 | 0.9338 | 0.9399 | 0.9405 |
| Resnet50Unet | 0.9142 | 0.9285 | 0.9289 | 0.9142 | 0.9285 | 0.9289 | 0.9143 | 0.9285 | 0.9289 |
| VggUnet | 0.9162 | 0.9127 | 0.9130 | 0.8952 | 0.9118 | 0.9181 | 0.8952 | 0.9084 | 0.9147 |

Table 6 – Cumulative averaging

| Score | Architecture code name | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | G | H | I | J |
| DSC | 0.9193 | 0.9490 | 0.9379 | 0.9088 | 0.9348 | 0.9504 | 0.9054 | 0.9324 | 0.9296 | 0.9164 |

Table 7 – Fcn8Mobilenet trained on 5 different datasets

| Score | Network number | | | | | Statistics | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Min | Max | Avg |
| DSC | 0.9004 | 0.8441 | 0.8607 | 0.8881 | 0.8773 | 0.8441 | 0.9004 | 0.8741 |

Table 8 – The result of averaging Fcn8Mobilenet trained on 5 datasets

| Score | Number of networks in the ensemble | | | | Statistics | | |
|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | Min | Max | Avg |
| DSC | 0.9023 | 0.8789 | 0.9023 | 0.9006 | 0.8789 | 0.9023 | 0.8960 |

Table 9 – Result of averaging a large number of Fcn8Mobilenet in one ensemble

| Score | Number of networks in the ensemble | | | | | | | | | | Statistics | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | Min | Max | Avg |
| DSC | 0.9240 | 0.9252 | 0.9242 | 0.9251 | 0.9254 | 0.9259 | 0.9257 | 0.9250 | 0.9255 | 0.9261 | 0.9185 | 0.9261 | 0.9246 |

Table 10 – The result of applying the mask shape averaging method

| Architecture name | Number of networks in the ensemble | | | | Statistics | | |
|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | Min | Max | Avg |
| Fcn8Mobilenet | 0.925156 | 0.92502 | 0.920583 | 0.922195 | 0.920583 | 0.925156 | 0.923239 |
| Fcn8Resnet50 | 0.951575 | 0.945714 | 0.944762 | 0.944422 | 0.944422 | 0.951575 | 0.946618 |
| Fcn8Vgg | 0.938029 | 0.930131 | 0.93025 | 0.927827 | 0.927827 | 0.938029 | 0.931559 |
| MobilenetUnet | 0.913377 | 0.91427 | 0.910609 | 0.906647 | 0.906647 | 0.91427 | 0.911226 |
| Pspnet50 | 0.934716 | 0.933914 | 0.935697 | 0.931073 | 0.931073 | 0.935697 | 0.93385 |
| Pspnet101 | 0.94993 | 0.944895 | 0.93448 | 0.856578 | 0.856578 | 0.94993 | 0.921471 |
| Resnet50Pspnet | 0.905457 | 0.909698 | 0.912497 | 0.906793 | 0.905457 | 0.912497 | 0.908611 |
| Resnet50Segnet | 0.935092 | 0.93356 | 0.932842 | 0.92569 | 0.92569 | 0.935092 | 0.931796 |
| Resnet50Unet | 0.937276 | 0.937114 | 0.931678 | 0.929995 | 0.929995 | 0.937276 | 0.934016 |
| VggUnet | 0.919379 | 0.915692 | 0.912555 | 0.916767 | 0.912555 | 0.919379 | 0.916098 |

## 6 DISCUSSION

Having considered the comparison of different pre-processing methods, we can conclude that the use of such methods is undesirable for the tasks of segmentation of chest X-ray images. The use of these methods did not produce the desired result, but on the contrary, made it worse. The use of these methods reduced the visibility of boundaries, blurred the image, and some methods even increased noise. The presence of these methods in the segmentation algorithm complicates the extraction of the necessary features from images and is definitely not recommended.

The use of averaging method is the simplest ensemble method, which makes it easy to use. However as the result showed, only in 12 cases out of 40 could an improvement be achieved compared to using single networks. This method, in the presence of a large number of predictions of poor quality, has the property of deteriorating the resulting prediction. The use of this method in problems that require guaranteed accuracy is ambiguous and undesirable. At the same time, the study of this method proved that the use of ensembles of different architectures has an advantage over the use of ensembles of the same architectures due to the combination of different data extraction methods.


DOI 10.15588/1607-3274-2024-4-10

The weighted averaging method was able to show improvement over single networks in 22 cases out of 90, suggesting that applying weights to networks would strengthen strong networks and weaken weak networks. However, this method did not show superiority over the averaging method. As can be seen in Table 5, the weights for this method were determined based on three metrics. The result of this algorithm directly depends on the selection of the necessary weights, so further in-depth study of this method can improve its accuracy.

The cumulative method, based on the idea of averaging already averaged predictions, was expected to be better than averaging method. But this is only guaranteed if averaging is able to provide only positive results. Only in five out of ten cases compared to single networks and in two out of ten compared to ensembles of averaging did this method show an increase in results. Such results indicate that the use of averaging is unstable and further highlight the controversy over the use of averaging.

Using an ensemble of networks of the same architecture trained on different datasets helps to reduce the correlation of model errors. In three out of four cases, using ensembles with the usual averaging method was better than using a single network with better accuracy. Although the obtained result was small, it should not be forgotten that the used dataset also consists of augmented data and if the opportunity to use a dataset with a large number of unique images and its division into subsets was provided, the result could have been significantly better.

Using 100 trained networks showed the same ambiguous result. Although the ensemble of 100 networks was the most successful in comparison with other numbers of networks in the ensemble, the total weight of 252 GB of all networks is hard to imagine in use in a medical institution.

The method of averaging mask shapes using distance transformation was able to demonstrate a good result. This method was better in 20 cases compared to single networks, while conventional averaging brought a result of 12. It managed to show a result better than averaging, because the algorithm is based not on averaging pixel values, but on calculating mask distances. But this method has one nuance that stops it on the way to a complete improvement of the results. This method is poorly susceptible to artefacts in the form of growths on the mask, since the shapes of masks are averaged, any mask used whose shape will differ significantly from the shape of the desired object crosses out all the positive aspects of its work. To eliminate this interference, it is worth using methods for removing artifacts on masks which will definitely lead to an improvement in the work of this algorithm.

## CONCLUSIONS

In this study, different methods of combining neural network ensemble predictions for segmentation of the thoracic spine region were examined. The extensive study of different methods allowed us to further explore the choice and advantage of specific methods. In any case, ensemble methods were able to demonstrate improvement in segmentation, although not in all cases considered.

The impact of image preprocessing on X-ray image segmentation tasks was also studied. The obtained results gave reason to doubt the appropriateness of these methods for solving such problems.

**The scientific novelty** of the obtained results is the effect of such neural network prediction fusion algorithms using such neural network architectures was examined for the first time, and the effect of six image preprocessing methods for segmentation was studied. This allows us to select the appropriate method for further spine segmentation studies.

**The practical significance** consists in the fact that a comparison of several methods for combining mask predictions on chest X-ray images in the lateral projection was made, which made it possible to apply this approach to creating automatic segmentation of vertebrae or necessary areas of the spine and implement them in medical institutions.

**Prospects for further research** is a more detailed study of the weighted averaging method with a more extensive selection of weights that may depend on a large number of metrics or other parameters. Also, further improvement of the mask shape averaging method may bring a more successful result than was obtained.

## ACKNOWLEDGEMENTS

We express our gratitude to A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine for the opportunity to conduct this study.

## REFERENCES

1. Wu H., Wu X., Wu T., Miao X., Zheng S., Huang G., Cheng X. Detection Ewingella americana from a patient with Andersson lesion in ankylosing spondylitis by metagenomic next-generation sequencing test: a case report, *BMC Musculoskelet Disord*, 2024, Vol. 25, P. 568. DOI: 10.1186/s12891-024-07680-y
2. Zhou Y., Huang X., Liu Y., Zhou X., Liu Q. Destructive Cryptococcal Osteomyelitis Mimicking Tuberculous Spondylitis, *American Journal of Case Reports*, 2024, Vol. 25, P. e944291. DOI: 10.12659/AJCR.944291
3. Smolle M. A., Maier A., Lindenmann J., Porubsky C., Leithner J., Smolle-Juettner F. M. Esophageal perforation with near fatal mediastinitis secondary to Th3 fracture, *Wiener klinische Wochenschrift*, 2024. DOI: 10.1007/s00508-024-02397-3
4. Sønderby A. H., Thomsen H., Skals R. G., Storm S., Leutscher P.D.C., Simony A. Thoracic spine X-ray examination of patients with back pain using different breathing technique and exposure times – A diagnostic study, *Radiography*, 2024, Vol. 30, pp 582–288. DOI: 10.1016/j.radi.2024.01.011
5. Kjelle E., Chilanga C. The assessment of image quality and diagnostic value in X-ray images: a survey on radiographers' reasons for rejecting images, *Insights into imaging*, 2022, Vol. 13(1), № 36. DOI: 10.1186/s13244-022-01169-9
6. Ullman G. Quantifying image quality in diagnostic radiology using simulation of the imaging system and model observers. Linköping, Sweden, 2008, 85 p.
7. Dietterich T. G. Ensemble Methods in Machine Learning, *Multiple Classifier Systems. MCS 2000. Springer. Berlin,*


DOI 10.15588/1607-3274-2024-4-10

*Heidelberg*, 2000, pp. 1–15. (Lecture Notes in Computer Science, Vol. 1857). DOI: 10.1007/3-540-45014-9_1
8. Mohammed A., Kora R. A comprehensive review on ensemble deep learning: Opportunities and challenges, *Journal of King Saud University – Computer and Information Sciences*, 2023, Vol. 35(2), pp. 757–774. DOI: 10.1016/j.jksuci.2023.01.014
9. Maclin R., Opitz D. W. Popular Ensemble Methods: An Empirical Study, *Journal of Artificial Intelligence Research*, 1999, Vol. 11, pp. 169–198.
10. Lu H., Li M., Zhang Y., Yu L. Lumbar spine segmentation method based on deep learning, *Journal of applied clinical medical physics*, 2023, Vol. 24, Iss. 6, P. e13996. DOI: 10.1002/acm2.13996
11. Xiong X. Graves S. A., Gross B. A., Buatti J. M., Beichel R. R. Lumbar and Thoracic Vertebrae Segmentation in CT Scans Using a 3D Multi-Object Localization and Segmentation CNN, *Tomography*, 2024, Vol. 10(5), pp. 738–760. DOI: 10.3390/tomography10050057
12. Li H., Luo H., Huan W., Shi Z., Yan C., Wang L., Mu Y., Liu Y. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network, *Neural computing & applications*, 2021, Vol. 33(18), pp. 11589–11602. DOI: 10.1007/s00521-021-05856-4
13. Khandelwal P., Collins L. D., Siddiqi K. Spine and Individual Vertebrae Segmentation in Computed Tomography Images Using Geometric Flows and Shape Priors, *Frontiers in Computer Science*, 2021, Vol. 3. DOI: 10.3389/fcomp.2021.592296
14. Mushtaq M., Akram M. U., Alghamdi N. S., Fatima J., Masood R. F. Localization and Edge-Based Segmentation of Lumbar Spine Vertebrae to Identify the Deformities Using Deep Learning Models, *Sensors*, 2022, Vol. 22(4), P. 1547. DOI: 10.3390/s22041547
15. Liang Y., Fang Y. T., Lin T. C., Yang C. R., Chang C. C., Chang H. K., Ko C. C., Tu T. H., Fay L. Y., Wu J. C., Huang W. C., Hu H. W., Chen Y. Y., Kuo C. H. The Quantitative Evaluation of Automatic Segmentation in Lumbar Magnetic Resonance Images, *Neurospine*, 2024, Vol. 21(2). pp. 665–675. DOI: 10.14245/ns.2448060.030
16. Zhou Z., Wang S., Zhang S., Pan X., Yang H., Zhuang Y., Lu Z. Deep learning-based spinal canal segmentation of computed tomography image for disease diagnosis: A proposed system for spinal stenosis diagnosis, *Medicine*, 2024, Vol. 103(18), P. e37943. DOI: 10.1097/MD.0000000000037943
17. Vindr.ai Datasets: SpineXR. [Electronic resource]. Access mode: https://vindr.ai/datasets/spinexr

Received 02.09.2024.
Accepted 24.10.2024.

УДК 004.8

## ВПЛИВ ПОПЕРЕДНЬОЇ ОБРОБКИ ТА ПОРІВНЯННЯ НЕЙРОМЕРЕЖЕВИХ АНСАМБЛЕВИХ МЕТОДІВ ДЛЯ СЕГМЕНТАЦІЯ ГРУДНОГО ВІДДІЛУ ХРЕБТА НА РЕНТГЕНІВСЬКИХ ЗНІМКАХ

**Конюхов В. Д**. – аспірант, Інститут енергетичних машин і систем ім. А. М. Підгорного НАН України, Харків, Україна.
**Моргун О. М.** – канд. фіз.-мат. наук, директор ТОВ «Лабораторія рентгенівської медичної техніки», Харків, Україна.
**Нємченко К. Е.** – д-р фіз.-мат. наук, завідувач кафедри Харківський національний університет імені В. Н. Каразіна, Харків, Україна.

### АНОТАЦІЯ

**Актуальність.** Автоматична сегментація медичних знімків відіграє важливу роль у процесі автоматизації визначення захворювань різного роду області хребта, а використання рентгенографії є найдоступнішим засобом передбачення захворювань. За багато років було проведено безліч досліджень на тему сегментації зображень. Одним із багатьох методів покращення сегментації зображень є застосування ансамблів нейронних мереж.

**Метою** даного дослідження було розглянути вплив попередньої обробки зображень та вивчити і порівняти головні методи ансамблів нейронних мереж та їх вплив на сегментацію області хребта, в даному дослідженні розглядалася область яка складається з хребців: Th8, Th9, Th10, Th11.

**Метод.** Для початку було розглянуто вплив попередньої обробки рентгенівських зображень, яка включала в себе наступні методи: вирівнювання гістограми для поліпшення контрасту, адаптивне вирівнювання гістограми з обмеженням контрасту, метод логарифмічного перетворення, медіанний фільтр, Гауссово згладжування. Для вивчення впливу ансамблю нейронних мереж на якість сегментації використовувалися такі методи: метод усереднення – найпростіший метод половинного усереднення; зважене усереднення – покращена версія методу усереднення, яка використовує ваги для кожної мережі, чим більша вага мережі – тим більший її вплив на усереднення; метод усереднення усереднених зображень – модифікований метод усереднення в якому кожен ансамбль отримує усереднене зображення, після чого всі результати ансамблів усереднюються; метод усереднення мереж навчених на різних даних – використовується *n* мереж, навчальна вибірка розбивається на *n* частин, кожна нейронна мережа навчається на своїй підмножині даних, в результаті для передбачень використовується звичайний метод усереднення; метод усереднення для великої кількості мереж – у цьому методі було навчено 100 нейронних мереж, після чого використовувався звичайний метод усереднення; метод усереднення контурів – даний метод усереднює всі контури в результаті чого виходить один середній контур.

**Результати.** Було досліджено, що застосування різних методів попередньої обробки зображень не гарантує поліпшення якості сегментації області хребта на рентгенівських знімках, а навіть навпаки погіршує якість сегментації. Були розглянуті різні методи об’єднання передбачень ансамблів нейронних мереж, що дало можливість дізнатися плюси та мінуси конкретних методів для завдання сегментації рентгенівських знімків.

**Висновки.** Проведені експерименти дали можливість зробити висновок, що застосування будь-яких методів попередньої обробки не варто використовувати для сегментації рентгенівських знімків. Також завдяки великій кількості архітектур і методів об’єднання передбачень було вивчено поведінку ансамблевих методів що дозволить надалі визначити необхідний підхід для сегментації рентгенівських знімків. Подальше вивчення методу зваженого усереднення і методу усереднення форм масок дасть можливість поліпшити отриманий результат і досягти ще більшого успіху в сегментації.

**КЛЮЧОВІ СЛОВА:** машинне навчання, розпізнавання образів, нейронна мережа, сегментація зображення, комп’ютерний зір.


DOI 10.15588/1607-3274-2024-4-10

**ЛІТЕРАТУРА**

1. Detection Ewingella americana from a patient with Andersson lesion in ankylosing spondylitis by metagenomic next-generation sequencing test: a case report / [H. Wu, X. Wu, T. Wu et al.] // BMC Musculoskelet Disord. – 2024. – Vol. 25. – P. 568. DOI: 10.1186/s12891-024-07680-y
2. Destructive Cryptococcal Osteomyelitis Mimicking Tuberculous Spondylitis / [Y. Zhou, X. Huang, Y. Liu et al.] // American Journal of Case Reports. – 2024. – Vol. 25. – P. e944291. DOI: 10.12659/AJCR.944291
3. Esophageal perforation with near fatal mediastinitis secondary to Th3 fracture / [M. A. Smolle, A. Maier, J. Lindenmann et al.] // Wiener klinische Wochenschrift. – 2024. DOI: 10.1007/s00508-024-02397-3
4. Thoracic spine X-ray examination of patients with back pain using different breathing technique and exposure times – A diagnostic study / [A. H. Sønderby, H. Thomsen, R. G. Skals et al.] // Radiography. – 2024. – Vol. 30. – P. 582–288. DOI: 10.1016/j.radi.2024.01.011
5. Kjelle E. The assessment of image quality and diagnostic value in X-ray images: a survey on radiographers' reasons for rejecting images / E. Kjelle, C. Chilanga // Insights into imaging. – 2022. – Vol. 13(1), № 36. DOI: 10.1186/s13244-022-01169-9
6. Ullman G. Quantifying image quality in diagnostic radiology using simulation of the imaging system and model observers / G. Ullman. – Linköping, Sweden, 2008. – 85 p.
7. Dietterich T. G. Ensemble Methods in Machine Learning / T. G. Dietterich // Multiple Classifier Systems. MCS 2000. Springer. – Berlin, Heidelberg, 2000. – P. 1–15. – (Lecture Notes in Computer Science, Vol. 1857). DOI: 10.1007/3-540-45014-9_1
8. Mohammed A. A comprehensive review on ensemble deep learning: Opportunities and challenges / A. Mohammed, R. Kora // Journal of King Saud University – Computer and Information Sciences. – 2023. – Vol. 35(2). – P. 757–774. DOI: 10.1016/j.jksuci.2023.01.014
9. Maclin R. Popular Ensemble Methods: An Empirical Study / R. Maclin, D. W. Opitz // Journal of Artificial Intelligence Research. – 1999. – Vol. 11. – P. 169–198.
10. Lumbar spine segmentation method based on deep learning / [H. Lu, M. Li, Y. Zhang, L. Yu] // Journal of applied clinical medical physics. – 2023. – Vol. 24(6). – P. e13996. DOI: 10.1002/acm2.13996
11. Xiong X. Lumbar and Thoracic Vertebrae Segmentation in CT Scans Using a 3D Multi-Object Localization and Segmentation CNN / [X. Xiong, S. A. Graves, B. A. Gross et al.] // Tomography. – 2024. – Vol. 10(5). – P. 738–760. DOI: 10.3390/tomography10050057
12. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network / [H. Li, H. Luo, W. Huan et al.] // Neural computing & applications. – 2021. – Vol. 33(18). – P. 11589–11602. DOI: 10.1007/s00521-021-05856-4
13. Khandelwal P. Spine and Individual Vertebrae Segmentation in Computed Tomography Images Using Geometric Flows and Shape Priors / P. Khandelwal, L. D. Collins, K. Siddiqi // Frontiers in Computer Science. – 2021. – Vol. 3. DOI: 10.3389/fcomp.2021.592296
14. Localization and Edge-Based Segmentation of Lumbar Spine Vertebrae to Identify the Deformities Using Deep Learning Models / [M. Mushtaq, M. U. Akram, N. S. Alghamdi et al.] // Sensors. – 2022. – Vol. 22(4). – P. 1547. DOI: 10.3390/s22041547
15. The Quantitative Evaluation of Automatic Segmentation in Lumbar Magnetic Resonance Images / [Y. W. Liang, Y. T. Fang, T. C. Lin et al.] // Neurospine. – 2024. – Vol. 21(2). – P. 665–675. DOI: 10.14245/ns.2448060.030
16. Zhou, Z. Deep learning-based spinal canal segmentation of computed tomography image for disease diagnosis: A proposed system for spinal stenosis diagnosis / [Z. Zhou, S. Wang, S. Zhang et al.] // Medicine. – 2024. – Vol. 103(18). – P. e37943. DOI: 10.1097/MD.0000000000037943
17. Vindr.ai Datasets: SpineXR. [Electronic resource]. – Access mode: https://vindr.ai/datasets/spinexr


DOI 10.15588/1607-3274-2024-4-10

p-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2024. № 4
e-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2024. № 4

UDC 004.8

# ENSEMBLE METHOD BASED ON AVERAGING SHAPES OF OBJECTS USING THE PYRAMID METHOD

**Koniukhov V. D.** – Postgraduate student, A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine, Kharkiv, Ukraine.

## ABSTRACT

**Context.** Image segmentation plays a key role in computer vision. The quality of segmentation is affected by many factors: noise, artifacts, complex shapes of objects. Classical methods cannot always guarantee good success, depending on the quality of the image and the existing noise, they cannot always achieve the desired result. The proposed method uses an ensemble of neural networks, which makes it possible to increase the accuracy and stability of segmentation.

**Objective.** The goal of the work is to develop a new method of combining predictions of neural network ensembles, which can improve segmentation accuracy by combining images of different image sizes.

**Method.** A method is proposed that averages the shapes of objects depicted on prediction masks. A pyramid of images is used to improve segmentation quality, each level of the pyramid corresponds to an increased size of the original image. This approach allows obtaining image characteristics at different levels. For a test image, a prediction is obtained from each neural network in the ensemble, after which a pyramid is built for the image. All pyramid levels are combined into the final image using SAAMC. All obtained final images for each neural network are also combined at the end using SAAMC. The use of an ensemble of neural networks combined with the pyramid method allows for reducing the impact of noise and artifacts on the segmentation results.

**Results.** The use of this method was compared with the usual use of individual neural networks and the ensemble averaging method. The obtained results show that the proposed method outperforms its competitors. Application of the proposed method improved the accuracy and quality of segmentation.

**Conclusions.** The conducted research confirmed the sense of using an ensemble of neural networks and creating a new method of combining predictions. The use of an ensemble of neural networks makes it possible to compensate for the errors and shortcomings of individual neural networks. Using the proposed method can significantly reduce the impact of noise and artifacts on segmentation. Further study and modification of this method will make it possible to further improve the quality of segmentation.

**KEYWORDS:** machine learning; image recognition; neural network; image segmentation, computer vision.

## ABBREVIATIONS

DSC is a Dice-Sorensen coefficient;
DCNN is a Deep Convolutional Neural Network;
FCN is a Fully Convolutional Network;
FPN is a Feature Pyramid Network;
SAAMC is a Shape Averaging with Alignment to a Mean Center.

## NOMENCLATURE

$A$ is a set of pixels of ground truth mask;
$\text{Accuracy}_{\text{new}}$ is an accuracy of new method;
$\text{Accuracy}_{\text{competitor}}$ is an accuracy of competitive method;
$B$ is a set of pixels of a predicted mask;
$C_i$ is $i$-th center of mass;
$C_{\text{mean}}$ is a mean center of mass;
$d$ is a difference between two distance maps;
$d_{\text{all}}$ is a difference between all distance maps;
$dt()$ is a distance transformation function;
$f(x)$ is a proposed method;
$i$ is an index of an element;
$I_i$ is $i$-th image;
$I_{\text{res}}$ is a resulting image;
$I_{\text{resized}}$ is a resized image;
$k$ is a pyramid level number;
$\text{mask}_i$ is $i$-th image mask;
~mask is an inverted image mask;
$M$ is a length of one side of an image;
$n$ is a number of elements;
$N$ is a number of pyramid levels;
$O$ is a length of one side of the input image;
$P_i$ is an $i$-th prediction;
$R_i$ is an image size at level $i$;
$S$ is a set of images;
$x_{\text{mean}}$ is a mean value of $x$ coordinates;
$y_{\text{mean}}$ is a mean value of $y$ coordinates.

## INTRODUCTION

Modern methods of image segmentation are mostly based on neural networks. The use of neural networks for segmentation tasks of various types of images has proven its effectiveness compared to classical methods [1]. Despite the achievements brought by the use of neural networks, there were still some difficulties that had to be faced: noise, artifacts, complex objects, instability of results. For this, new methods were developed, as well as new neural networks that could minimize the impact of various undesirable conditions on the segmentation result.

One of the improved approaches to solving such problems was the U-Net neural network [2], which had significant success in segmentation of medical images. But neural networks still continued to face significant challenges. For this, ensemble methods were proposed. This approach included combining the results of several neural networks and improving the quality of segmentation. Thanks to ensemble methods, it became possible to increase the accuracy and reliability of segmentation. Ensemble methods and image segmentation were not ignored, the proposed methods: averaging, weighted averaging, training one network on different data demon-


DOI 10.15588/1607-3274-2024-4-11

strated their effectiveness compared to previously created methods. The main role was played by the methods of unification. The ease of their use is one of the advantages. This article proposes to consider a new method that combines the advantage of combining masks and using different image sizes to extract image features at different levels. This method is aimed at increasing the accuracy of segmentation and smoothing out the bad influence of low-quality images.

**The object of study** is performance of the proposed method. The main focus is on combining the predictions of different neural networks, as well as using a different number of pyramid levels to extract different image characteristics. This study includes an analysis of the methods used and an assessment of their effectiveness on two data sets.

**The subject of study** is ensemble methods and the pyramid method. Existing methods have their pros and cons. The application of these methods to different data sets may have different results. The development of a new method that can improve the accuracy of segmentation and minimize the impact of existing noise is a necessity.

**The purpose of the work** is to develop a new ensemble method that can improve the accuracy of image segmentation and test its effectiveness.

## 1 PROBLEM STATEMENT

Image segmentation is a complex task in computer vision. Segmentation problems are especially relevant when using this approach in complex areas such as medicine, self-driving cars, robotics. The use of segmentation in such areas requires high reliability and accuracy of results, which traditional methods cannot cope with. One of the most difficult areas in which it is necessary to guarantee maximum reliability and accuracy of segmentation is the segmentation of medical images. The presence of noise, artifacts, incorrect exposure or the influence of poor-quality equipment creates significant difficulties in performing successful segmentation. Let us assume that a set of images of unknown quality is given:

$$S = \left\{ I_i \right\}_{i=1}^{n}.$$

For a given data set, the problem of developing a new ensemble method can be represented as:

$$\text{Accuracy}_{\text{new}} = f\left(S\right) \to \max.$$

To solve the problems associated with poor segmentation quality, it is necessary that the proposed method be better than its competitors:

$$\text{Accuracy}_{\text{new}} > \text{Accuracy}_{\text{competitor}}.$$

## 2 REVIEW OF THE LITERATURE

Ensemble methods have proven themselves to be highly effective methods for solving complex problems. The main idea of the ensemble approach is to combine several neural networks, which makes it possible to smooth out the errors of individual networks. Thanks to the use of ensembles, it became possible to combine different architectures, which in turn makes it possible to take into account different aspects and extract the necessary characteristics in different ways. Different neural networks can produce results differently, depending on the data used. Combining them makes it possible to minimize the receipt of erroneous predictions and get a better result.

The earliest ensemble methods were: bagging [3], boosting [4] and stacking [5]. Bagging allows training several models on different subsets of data, and at the end either voting for the final result or averaging. Boosting is based on the idea of improving each subsequent model. This method is used when a neural network gets bad results. Each neural network in the ensemble learns from the errors of the previous one. In stacking, several models are trained, after which their predictions are passed to a meta-model, which forms the final prediction based on the data.

In their paper [6], the authors present an analysis of published works on ensemble learning and ensemble deep learning. A new approach for retinal vessel segmentation based on deep ensemble learning is presented in the study [7]. The obtained results showed that the proposed model outperforms existing methods on different datasets. In the paper [8], a deep ensemble model was proposed for classification of histopathological images of the breast. The proposed method was compared with others, as a result of which it was able to provide an advantage of 5-20%. Ensemble learning for brain tumor classification was proposed in the paper [9]. As a preprocessing of the data, the authors used Gabor filters to remove noise. They used three models: EfficientNet, DenseNet and MobileNet in order to achieve diversity in feature extraction. The method proposed by the authors showed good results and promising performance. In the paper [10], the authors proposed a method for using ensemble learning to detect skin cancer. The authors argue that the use of individual models is not as reliable as desired and they cannot achieve the required accuracy. In turn, thanks to the use of an ensemble of neural networks, it was possible to achieve a better result.

A three-phase approach for sclera segmentation in eye images is presented in the paper [11]. In their study, the authors use five deep learning models: Unet-DenseNet, FPN with DenseNet, Unet-ResNet50, TransUnet, and Swin-UNet. Using such models in an ensemble ensures a different approach to extracting the necessary features and improving the quality of segmentation. With this approach, the authors minimized the impact of errors and were able to achieve a good result. The paper [12] discusses a heterogeneous ensemble approach based on DCNN. In their work, the authors use weighted averag-


DOI 10.15588/1607-3274-2024-4-11

ing. They assign larger weights to the models that showed the best result during testing. This approach helped to generalize the distribution and prove the advantage of using the ensemble method.

In the paper [13], the authors propose a pyramid attention method for image restoration. This method captures features that are located far from each other from a multi-scale pyramid of features. The proposed approach can be easily integrated into different neural architectures. A multi-scale feature pyramid fusion network is proposed in the paper [14]. The authors conducted the study on three datasets and compared their method with U-Net. Ultimately, the authors were able to achieve better results than the competitive network. Spectral-Spatial Feature Pyramid Network was proposed in [15]. This network extracts multi-scale spectral information and multi-scale spatial information using the attention mechanism and feature pyramid structure. The obtained results indicate that the network can achieve good results. In the paper [16], the authors propose a context-aware network with a two-stream pyramid (CANet). They use this network to segment medical images on three datasets. Using this approach gave a positive result compared to 13 competitive methods. By using a pyramid of multi-scale features [17], the authors were able to demonstrate the effectiveness of the pyramidal multi-scale structure for segmentation of histopathological images. The authors focus on adapting to changes in image resolution, which helps to improve the quality of segmentation.

## 3 MATERIALS AND METHODS

At the beginning of the construction of the proposed algorithm, $n$ number of neural networks is selected, which will comprise the ensemble.

First stage. Let there be an $I_i$ image of size M×M, in grayscale. It is necessary to build a pyramid of levels for it. This means building an image pyramid that will accept an image of size from the data set as input, will contain $n$ levels, where at each level the size of the original image will increase in size. In this study, the following image sizes were used in the pyramid: 64×64, 128×128, 256×256, 512×512, 1024×1024, 2048×2048 and 4096×4096. The pyramid can be designated as:

$$\text{Pyramid} = \{I_1, I_2 \dots I_n\},$$

where at each level of the pyramid the image is designated as:

$$I_k^{R_i}.$$

The image size was determined as follows:

$$R_i = \left\{\left(O \times 2^{i-1}\right) \times \left(O \times 2^{i-1}\right) | i \in [1,\dots, N]\right\}.$$

Second stage. After the pyramid has been built, it is necessary to obtain predictions for each level of the pyramid. The size of the prediction mask is equal to the size of the image for which the prediction was made.

Third stage. Here it is necessary to bring each level of the pyramid to a single size:

$$\forall i \in \{1,2\dots N\}: \text{Resize}(I_i) = I_{\text{resized},i}^{M \times M}.$$

After which all the images that are in the pyramid will be presented in the same size.

Fourth stage. The method of combination of predictions must be applied to all the obtained images at the third stage, in this case SAAMC will be used. First, it is necessary to find the average value of the center for all the predictions used. Let $n$ predictions be given $\{P_1, P_2 \dots P_n\}$, each object on the mask has a center of mass $\{C_1, C_2 \dots C_n\}$. The average value of the center for all objects will be $C_{\text{mean}} = (x_{\text{mean}}. y_{\text{mean}})$, where:

$$x_{\text{mean}} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and

$$y_{\text{mean}} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

After the average center has been obtained, it is necessary for each prediction $P_i$ to shift the object relative to its new center – the mean center. When all objects have been shifted relative to the mean center, it is necessary to average the shapes of the objects using the formula:

$$d_{\text{all}} = \sum_{i=1}^{n} d(\text{mask}_i),$$

where $d(mask)$ is the function given below:

$$d = dt(\text{mask}) - dt(\sim \text{mask}).$$

The distance transformation method calculates the minimum distance from the object pixel to the background pixel. The final averaging of the object shapes is as follows:

$$I_{\text{res}} = d_{\text{all}} > 0.$$

To average the shapes of objects, it is necessary to sum up the obtained results and if the sum is greater than 0, then the pixels become white, otherwise black.

Fifth stage. For each neural network in the ensemble, the final result was obtained at the fourth stage. At this stage, it is necessary to apply SAAMC, which was described in the fourth stage, to all images obtained at the fourth stage. This stage is similar to the fourth, only the input images need to be changed.


DOI 10.15588/1607-3274-2024-4-11

To summarize all of the above, we can outline the proposed method in the form of a list of tasks:

1. Create a pyramid.
2. Get predictions for all levels of the pyramid.
3. Resize all images resulting from the second step.
4. Perform averaging of objects depicted in the predictions obtained in the third stage.
5. Perform averaging of objects depicted in the predictions obtained in the fourth stage.

Visually, all stages of the proposed method can be seen in Figure 2, the stage number is indicated in the circle.

## 4 EXPERIMENTS

Data obtained from open sources were used to conduct the experiments [18, 19]. These were X-rays and lung masks for them. In order to conduct the study in more detail, it was decided to split this data set into two. As a result, two sets were obtained: one contains masks of the left lung, the second masks of the right lung. Each dataset contained 703 images, where 630 images were used for training and the remaining 73 were used for testing. All images were presented in 512×512 size and in grayscale. An example of the images used can be seen in Figure 1. On the left in the image is the mask of the right lung, on the right is the mask of the left lung.

The neural network used was FCN8-MobileNet. It combines the FCN architecture and the MobileNet base model. This network combines two architectures for more efficient feature extraction and improved segmentation. Its main advantage is also its lightweight architecture, which helps reduce the number of adjustable parameters and increase its speed.

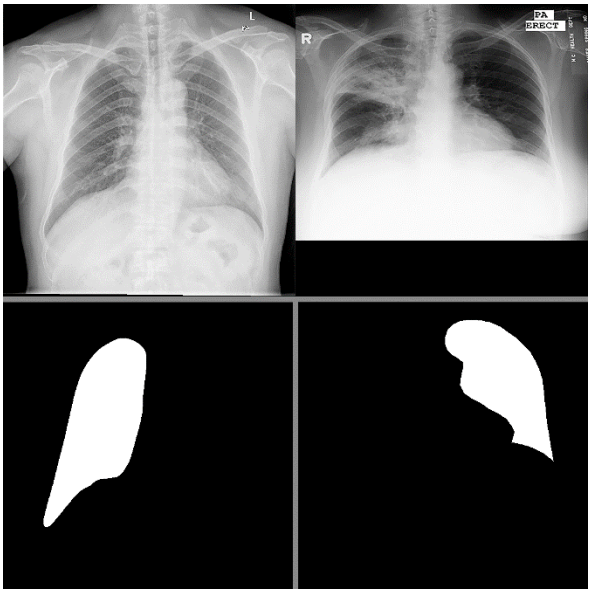



Figure 1 – The example of the images used and their masks

To check the similarity of the prediction and the ground truth mask, the Dice-Sorensen coefficient was used. It was represented by the following formula:

$$DSC = \frac{2|A \cap B|}{|A| + |B|}.$$

First of all, it was necessary to study the behavior of neural networks separately. For this purpose, 10 neural networks were trained on each data set. The next step was to obtain the DSC for different input image sizes. For this, one neural network was used, and the same operations were performed for each data set. An image resizing operation was applied to each test image. The following dimensions were used for this purpose: 64×64, 128×128, 256×256, 512×512, 1024×1024, 2048×2048 and 4096×4096. For each image with different sizes, a prediction was obtained. Then, for all images, the average DSC value was obtained and written down in a table.

The following experiment was used to justify the use of the ensemble. For this, an ensemble of 10 networks was used. Two methods were used to combine predictions in the ensemble: the well-known averaging method and SAAMC.

The use of the pyramid method for neural networks separately was also considered. For the combination of images obtained at each level of the pyramid, the same two methods were used: the averaging method and SAAMC. Table 1 shows the sizes of the images that were used to create the pyramids. The cross indicates the sizes of the images at each level of the pyramid. The images used are square, so only one side will be listed in the table.

Table 1 – Image sizes used at pyramid levels

| Name | Image sizes (M×M) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 64 | 128 | 256 | 512 | 1024 | 2048 | 4096 |
| A | X | X | X | | | | |
| B | | X | X | X | | | |
| C | | | | X | X | | |
| D | | | | X | X | X | |
| E | | | | X | X | X | X |
| F | X | X | X | X | X | X | X |

And at the very end, the use of an ensemble of neural networks for the pyramid method and methods of combining predictions was considered.

## 5 RESULTS

The DSC for all 20 trained neural networks is presented in Table 2. The obtained results indicate a good degree of segmentation. The small spread between neural networks indicates the stable operation of the model used.

Table 3 clearly shows how image size affects segmentation accuracy. The datasets used images of 512×512 pixels. Comparing this size with others, we can see that for this model, reducing the image size only worsened the situation. Increasing the image size, on the contrary, increased the accuracy of the prediction. For both sets, the best result was using images of 4096×4096.

The advantage of the ensemble is shown in Table 4. The segmentation quality of the ensemble is higher than the average quality of the neural networks separately. For


DOI 10.15588/1607-3274-2024-4-11

both data sets, the ensembles showed a better result. And SAAMC showed better accuracy than its analogue.

The studies of the use of pyramids with different levels for single networks were presented in Tables 5 and 6. They considered two different methods of combining level images. In Table 5, for the left and right lungs, the best result was using a pyramid that contained levels with the following image sizes: 512×512, 1024×1024, 2048×2048, 4096×4096. The best result for the left lung was able to exceed the average value for single networks. For the right lung, the situation is worse, the result obtained was less than the average of single networks.

In Table 6, for the left lung, to achieve the best result, it was necessary to use a pyramid with the following levels: 512×512, 1024×1024, 2048×2048, 4096×4096. For the right lung, the following sizes helped to achieve a good result: 512×512, 1024×1024, 2048×2048. In both cases, the results obtained exceeded the average for single networks.

The results obtained using the proposed method are in Table 7. For the averaging method, to achieve better results, it was necessary to use a pyramid with the following levels: 512×512, 1024×1024, 2048×2048, 4096×4096.

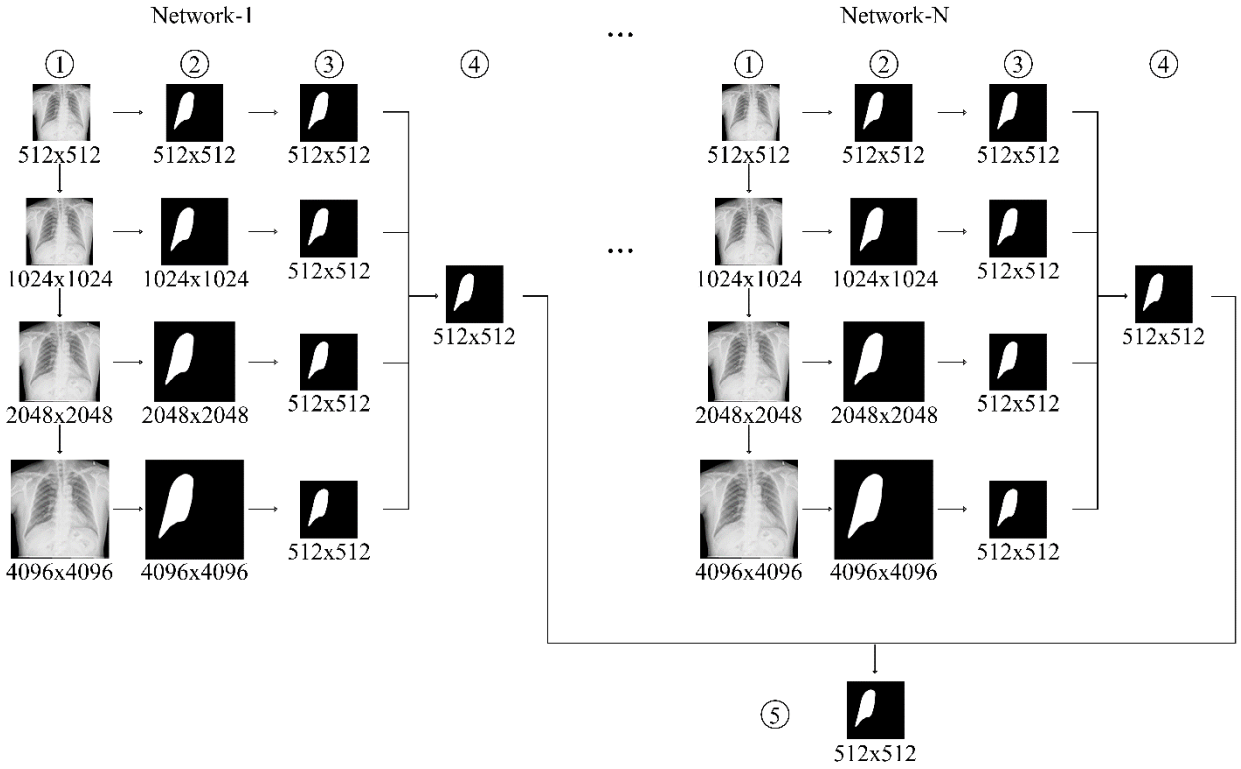



Figure 2 – Step-by-step example of the algorithm's operation.

Table 2 – Similarity measure for trained neural networks

| Dataset | Neural network number | | | | | | | | | | Statistics | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Min | Max | Mean |
| Left lung | 0.9540 | 0.9564 | 0.9559 | 0.9554 | 0.9553 | 0.9548 | 0.9575 | 0.9574 | 0.9560 | 0.9597 | 0.9540 | 0.9597 | 0.9562 |
| Right lung | 0.9485 | 0.9553 | 0.9519 | 0.9532 | 0.9488 | 0.9561 | 0.9580 | 0.9567 | 0.9568 | 0.9558 | 0.9485 | 0.9580 | 0.9541 |

Table 3 – Similarity measure for one neural network for different image sizes

| Dataset | Image size (M×M) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 64 | 128 | 256 | 512 | 1024 | 2048 | 4096 |
| Left lung | 0.8906 | 0.9356 | 0.9478 | 0.9540 | 0.9555 | 0.9563 | 0.9565 |
| Right lung | 0.8360 | 0.9219 | 0.9408 | 0.9485 | 0.9490 | 0.9499 | 0.9502 |

Table 4 – Similarity measure for an ensemble of neural networks

| Dataset | Combining method | |
|---|---|---|
| | Mean averaging | SAAMC |
| Left lung | 0.9599 | 0.9638 |
| Right lung | 0.9588 | 0.9610 |


DOI 10.15588/1607-3274-2024-4-11

Table 5 – Similarity measure using the pyramid method with mean averaging

| Dataset | Approach | Neural network number | | | | | | | | | | Statistics |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Mean |
| Left lung | A | 0.9210 | 0.9272 | 0.9289 | 0.9241 | 0.9211 | 0.9297 | 0.9357 | 0.9319 | 0.9269 | 0.9308 | 0.9277 |
| | B | 0.9418 | 0.9457 | 0.9471 | 0.9464 | 0.9417 | 0.9455 | 0.9499 | 0.9484 | 0.9454 | 0.9498 | 0.9462 |
| | C | 0.9542 | 0.9567 | 0.9559 | 0.9555 | 0.9555 | 0.9549 | 0.9576 | 0.9575 | 0.9563 | 0.9598 | 0.9564 |
| | D | 0.9543 | 0.9563 | 0.9574 | 0.9561 | 0.9544 | 0.9560 | 0.9588 | 0.9585 | 0.9561 | 0.9602 | 0.9568 |
| | E | 0.9557 | 0.9577 | 0.9585 | 0.9570 | 0.9560 | 0.9573 | 0.9600 | 0.9598 | 0.9573 | 0.9613 | 0.9580 |
| | F | 0.9539 | 0.9559 | 0.9572 | 0.9557 | 0.9540 | 0.9557 | 0.9586 | 0.9583 | 0.9557 | 0.9599 | 0.9565 |
| Right lung | A | 0.9007 | 0.9278 | 0.9137 | 0.9172 | 0.9154 | 0.9220 | 0.9274 | 0.9291 | 0.9312 | 0.9267 | 0.9007 |
| | B | 0.9335 | 0.9477 | 0.9377 | 0.9409 | 0.9370 | 0.9457 | 0.9501 | 0.9497 | 0.9485 | 0.9459 | 0.9335 |
| | C | 0.9487 | 0.9555 | 0.9520 | 0.9533 | 0.9489 | 0.9562 | 0.9581 | 0.9568 | 0.9569 | 0.9559 | 0.9487 |
| | D | 0.9479 | 0.9556 | 0.9515 | 0.9526 | 0.9489 | 0.9551 | 0.9578 | 0.9566 | 0.9569 | 0.9556 | 0.9479 |
| | E | 0.9492 | 0.9563 | 0.9525 | 0.9538 | 0.9499 | 0.9559 | 0.9583 | 0.9571 | 0.9576 | 0.9566 | 0.9492 |
| | F | 0.9472 | 0.9554 | 0.9509 | 0.9520 | 0.9485 | 0.9548 | 0.9575 | 0.9564 | 0.9566 | 0.9552 | 0.9472 |

Table 6 – Similarity measure using the pyramid method with SAAMC

| Dataset | Approach | Neural network number | | | | | | | | | | Statistics |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Mean |
| Left lung | A | 0.9424 | 0.9474 | 0.9362 | 0.9377 | 0.9498 | 0.9425 | 0.9456 | 0.9476 | 0.9447 | 0.9494 | 0.9443 |
| | B | 0.9539 | 0.9578 | 0.9529 | 0.9522 | 0.9574 | 0.9536 | 0.9563 | 0.9562 | 0.9552 | 0.958 | 0.9554 |
| | C | 0.9560 | 0.9571 | 0.9578 | 0.9562 | 0.9558 | 0.9562 | 0.9590 | 0.9589 | 0.9571 | 0.9605 | 0.9575 |
| | D | 0.9579 | 0.9591 | 0.9587 | 0.9575 | 0.9580 | 0.9580 | 0.9608 | 0.9606 | 0.9585 | 0.9622 | 0.9591 |
| | E | 0.9580 | 0.9590 | 0.9594 | 0.9577 | 0.9577 | 0.9584 | 0.9612 | 0.9609 | 0.9586 | 0.9624 | 0.9593 |
| | F | 0.9551 | 0.9580 | 0.9533 | 0.9518 | 0.9576 | 0.9550 | 0.9582 | 0.9580 | 0.9561 | 0.9600 | 0.9563 |
| Right lung | A | 0.9229 | 0.9342 | 0.9157 | 0.9312 | 0.9260 | 0.9415 | 0.9343 | 0.9335 | 0.9370 | 0.9345 | 0.9311 |
| | B | 0.9448 | 0.9515 | 0.9468 | 0.9501 | 0.9450 | 0.9548 | 0.9545 | 0.9528 | 0.9532 | 0.9525 | 0.9506 |
| | C | 0.9486 | 0.9561 | 0.9522 | 0.9533 | 0.9496 | 0.9562 | 0.9581 | 0.9570 | 0.9571 | 0.9561 | 0.9544 |
| | D | 0.9506 | 0.9568 | 0.9539 | 0.9550 | 0.9510 | 0.9572 | 0.9590 | 0.9577 | 0.9580 | 0.9575 | 0.9557 |
| | E | 0.9505 | 0.9568 | 0.9536 | 0.9548 | 0.9511 | 0.9569 | 0.9588 | 0.9578 | 0.9580 | 0.9575 | 0.9556 |
| | F | 0.9425 | 0.9510 | 0.9421 | 0.9492 | 0.9435 | 0.9548 | 0.9528 | 0.9515 | 0.9529 | 0.9530 | 0.9493 |

Table 7 – Similarity measure for an ensemble using the pyramid method with SAAMC

| Dataset | Approach | Method | |
|---|---|---|---|
| | | Mean averaging | SAAMC |
| Left lung | A | 0.9327 | 0.9524 |
| | B | 0.9505 | 0.9629 |
| | C | 0.9607 | 0.9646 |
| | D | 0.9615 | 0.9666 |
| | E | 0.9627 | 0.9666 |
| | F | 0.9611 | 0.9634 |
| Right lung | A | 0.9268 | 0.9396 |
| | B | 0.9485 | 0.9573 |
| | C | 0.9594 | 0.9612 |
| | D | 0.9592 | 0.9623 |
| | E | 0.9602 | 0.9622 |
| | F | 0.9588 | 0.9561 |

When applying the proposed method for the left lung, it was necessary to use a pyramid with the following level sizes: 512×512, 1024×1024, 2048×2048. For the right lung: 512×512, 1024×1024, 2048×2048, 4096×4096. As can be seen from the tables, in all cases SAAMC was better than average averaging. For the left lung, the improvement was 1.034, for the right 0.8236 compared to using single networks.

## 6 DISCUSSION

Starting from studying the simple use of ensembles, one can already notice the advantage of SAAMC over averaging. The study of the use of different sizes presented in Table 3 shows that the network that was used extracted features from large-size images better. In this case, the use of the pyramidal method made it possible to improve the quality of segmentation by extracting from different levels. This suggests that using higher resolution images allows the model to detect finer details.

By combining the results of 10 networks, we were able to prove the advantage of using the ensemble method. Using the ensemble allows us to smooth out the shortcomings of individual networks and enhance the final result. In this way, we can minimize the impact of noise, artifacts, and other undesirable moments.

The most important thing was the confirmation of the advantage of SAAMC over averaging and using single networks.

One of the main advantages of the method used is its ease of use and guaranteed results. Using FPN imposes a limitation in the form of increased computational costs. The higher the image resolution, the longer it takes to process, which means the model itself will take many times longer to learn. It is worth noting that the improvement in accuracy was noticed just when the resolution increased and larger images were used. This proves the main advantage of this method. To use it, there is no need to train a special neural network, as well as configure a


DOI 10.15588/1607-3274-2024-4-11

large number of parameters. All we need is to simply apply the proposed algorithm to an ensemble of neural networks. Naturally, we can combine different architectures, which can further improve the accuracy of the prediction. And using FPN as neural networks for an ensemble can also improve accuracy, but do not forget about the speed.

## CONCLUSIONS

A new ensemble method for combining predictions was proposed. The proposed method demonstrated its effectiveness in improving the quality of segmentation. Using the pyramid method as a basis for the developed method made it possible to extract features from different levels. Using the ensemble approach made it possible to smooth out the shortcomings of individual neural networks and combine their strengths. The main advantage of the method is its simplicity in implementation compared to the creation of FPN networks.

**The scientific novelty** of the obtained results is for the first time, a combination method was proposed that combines the advantages of the pyramid method and SAAMC. The results obtained confirmed the effectiveness of this method. The use of different resolutions at different levels of the pyramid made it possible to extract more detailed data, and the use of an ensemble of neural networks made it possible to combine the strengths of the neural networks used.

**The practical significance** of obtained results is that the use of the proposed method was tested on two medical data sets containing X-ray images. The use of this method can be easily implemented in automated systems for segmentation of medical images.

**Prospects for further research** are further developments of modifications of this method for further improvement of segmentation quality. For example, studies of using a large number of neural networks in ensembles, as well as using different architectures.

## ACKNOWLEDGEMENTS

I would like to express my gratitude to A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine for the opportunity to conduct this study.

## REFERENCES

1. Voulodimos A., Doulamis N., Doulamis A., Protopapadakis E. Deep Learning for Computer Vision: A Brief Review, *Computational intelligence and neuroscience*, 2018, P. 2018:7068349. DOI: 10.1155/2018/7068349
2. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, *ArXiv*, 2015. DOI: 10.1007/978-3-319-24574-4_28
3. Breiman L. Bagging predictors, *Machine Learning*, 1996, Vol. 24, No. 2, pp. 123–140. DOI: 10.1007/BF00058655
4. Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences*, 1997, Vol. 55, No. 1, pp. 119–139. DOI: 10.1006/jcss.1997.1504
5. Wolpert D. H. Stacked generalization, *Neural Networks*. – 1992, Vol. 5, No. 2, pp. 241–259. DOI: 10.1016/S0893-6080(05)80023-1
6. Yang Y., Lv H., Chen N. A survey on ensemble learning under the era of deep learning, *Artificial Intelligence Review*, 2023, Vol. 56, pp. 5545–5589. DOI: 10.1007/s10462-022-10283-5
7. Du L., Liu H., Zhang L., Lu Y., Li M., Hu Y., Zhang Y. Deep ensemble learning for accurate retinal vessel segmentation, *Computers in Biology and Medicine*, 2023, Vol. 158, P. 106829. DOI: 10.1016/j.compbiomed.2023.106829
8. Zheng Y., Li C., Zhou X., Chen H., Xu H., Li Y., Zhang H., Li X., Sun H., Huang X., Grzegorzek M. Application of transfer learning and ensemble learning in image-level classification for breast histopathology, *Intelligent Medicine*, 2023, Vol. 3, No. 2, pp. 115–128. DOI: 10.1016/j.imed.2022.05.004
9. Vaiyapuri J., Mahalingam S., Ahmad H. A. M., Abdeljaber E., Yang E., Jeong S. Y. Ensemble learning driven computer-aided diagnosis model for brain tumor classification on magnetic resonance imaging, *IEEE Access*, 2023, Vol. 11, pp. 91398–91406. DOI: 10.1109/ACCESS.2023.3306961
10. Tembhurne J. V., Hebbar N., Patil H. Y. et al. Skin cancer detection using ensemble of machine learning and deep learning techniques, *Multimedia Tools and Applications*, 2023, Vol. 82, pp. 27501–27524. DOI: 10.1007/s11042-023-14697-3
11. Al-Zebari A., Ensemble convolutional neural networks and transformer-based segmentation methods for achieving accurate sclera segmentation in eye images, *Signal, Image and Video Processing*, 2024, Vol. 18, pp. 1879–1891. DOI: 10.1007/s11760-023-02891-7
12. Iqball T., Wani M. A. Weighted ensemble model for image classification, *International Journal of Information Technology*, 2023, Vol. 15, pp. 557–564. DOI: 10.1007/s41870-022-01149-8
13. Mei Y., Fan Y., Zhang Y., Yu J., Zhou Y., Liu D., Fu Y., Huang T., Shi H. Pyramid attention network for image restoration, *International Journal of Computer Vision*, 2023, Vol. 131, pp. 3207–3225. DOI: 10.1007/s11263-023-01843-5
14. Zhang B., Wang Y., Ding C., Deng Z., Li L., Qin Z., Ding Z., Bian L., Yang C. Multi-scale feature pyramid fusion network for medical image segmentation, *International Journal of Computer Assisted Radiology and Surgery*, 2023, Vol. 18, pp. 353–365. DOI: 10.1007/s11548-022-02738-5
15. Fang L., Jiang Y., Yan Y., Yue J., Deng Y. Hyperspectral image instance segmentation using spectral-spatial feature pyramid network, *IEEE Transactions on Geoscience and Remote Sensing*, 2023, Vol. 61, pp. 1–13. DOI: 10.1109/TGRS.2023.3240481
16. Xie X., Zhang W., Pan X., Xie L., Shao F., Zhao W., An J. CANet: Context aware network with dual-stream pyramid for medical image segmentation, *Biomedical Signal Processing and Control*, 2023, Vol. 81, P. 104437. DOI: 10.1016/j.bspc.2022.104437
17. Qin P., Chen J., Zeng J., Chai R., Wang L. Large-scale tissue histopathology image segmentation based on feature pyramid, *EURASIP Journal on Image and Video Processing*, 2018. DOI: 10.1186/s13640-018-0320-8
18. Jaeger S., Karargyris A., Candemir S., Folio L., Siegelman J., Callaghan F., Xue Z., Palaniappan K., Singh R. K., Antani S., Thoma G., Wang Y. X., Lu P. X., McDonald C. J. Automatic tuberculosis screening using chest radiographs, *IEEE Transactions on Medical Imaging*, 2014, Vol. 33, No. 2, pp. 233–245. DOI: 10.1109/TMI.2013.2284099
19. Candemir S., Jaeger S., Palaniappan K., Musco J. P., Singh R. K., Xue Z., Karargyris A., Antani S., Thoma G., McDonald C. J. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration, *IEEE Transactions on Medical Imaging*, 2014, Vol. 33, No. 2, pp. 577–590. DOI: 10.1109/TMI.2013.2290491

Received 12.08.2024.
Accepted 15.10.2024.


DOI 10.15588/1607-3274-2024-4-11

УДК 004.8

# АНСАМБЛЕВИЙ МЕТОД ЗАСНОВАНИЙ НА УСЕРЕДНЕННІ ФОРМ ОБ'ЄКТІВ, ВИКОРИСТОВУЮЧИ ПІРАМІДНИЙ МЕТОД

**Конюхов В. Д**. – аспірант, Інститут енергетичних машин і систем ім. А. М. Підгорного НАН України, Харків, Україна.

## АНОТАЦІЯ

**Актуальність.** Сегментація зображень відіграє ключову роль в комп'ютерному зорі. На якість сегментації впливає багато факторів: шум, артефакти, складні форми об'єктів. Класичні методи не завжди можуть гарантувати гарний успіх, в залежності від якості зображення та наявного шуму, вони не завжди можуть досягти бажаного результату. Запропонований метод використовує ансамбль нейронних мереж, що дає змогу підвищити точність та стабільність сегментації.

**Мета роботи** – розробити новий метод комбінування передбачень ансамблю нейронних мереж, який зможе покращити точність сегментації за рахунок комбінування зображень різного розміру зображень.

**Метод.** Запропоновано метод який виконує усереднення форм об'єктів зображених на масках-передбаченнях. Для досягнення покращення якості сегментації використовується піраміда зображень, кожен рівень піраміди відповідає збільшеному розміру початкового зображення. Такий підхід дозволяє отримувати характеристики зображення на різних рівнях. Для тестового зображення отримується передбачення від кожної нейронної мережі в ансамблі, після чого для зображення будується піраміда. Всі рівні піраміди комбінуються в фінальне зображення за допомогою метода усереднення форм об'єктів. Всі отримані фінальні зображення для кожної нейронної мережі в кінці також комбінуються за допомогою метода усереднення форм об'єктів. Використання ансамблю нейронних мереж та пірамідного методу дають змогу зменшити вплив шумів та артефактів на результат сегментації.

**Результати.** Використання даного методу було порівняно зі звичайним використанням окремих нейронних мереж та ансамблевим методом усереднення. Отримані результати показують, що запропонований метод перевершує своїх конкурентів. Застосування запропонованого методу покращило точність та якість сегментації.

**Висновки.** Проведене дослідження підтвердило сенс використання ансамблю нейронних мереж та створення нового методу комбінування передбачень. Використання ансамблю нейронних мереж дає можливість компенсувати помилки та недоліки окремих нейронних мереж. Використання запропонованого методу може знизити вплив шумів та артефактів на сегментацію. Подальше вивчення та модифікація цього методу дадуть змогу покращити ще більше якість сегментації.

**КЛЮЧОВІ СЛОВА:** машинне навчання, розпізнавання образів, нейронна мережа, сегментація зображення, комп'ютерний зір.

## ЛІТЕРАТУРА

1. Deep Learning for Computer Vision: A Brief Review / [A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis] // Computational intelligence and neuroscience. – 2018. – P. 2018:7068349. DOI: 10.1155/2018/7068349
2. Ronneberger O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // ArXiv. – 2015. DOI: 10.1007/978-3-319-24574-4_28
3. Breiman L. Bagging predictors / L. Breiman // Machine Learning. – 1996. – Vol. 24, No. 2. – P. 123–140. DOI: 10.1007/BF00058655
4. Freund Y. A decision-theoretic generalization of on-line learning and an application to boosting / Y. Freund, R. E. Schapire // Journal of Computer and System Sciences. – 1997. – Vol. 55, No. 1. – P. 119–139. DOI: 10.1006/jcss.1997.1504
5. Wolpert, D. H. Stacked generalization / D. H. Wolpert // Neural Networks. – 1992. – Vol. 5, No. 2. – P. 241–259. DOI: 10.1016/S0893-6080(05)80023-1
6. Yang Y. A survey on ensemble learning under the era of deep learning / Y. Yang, H. Lv, N. Chen // Artificial Intelligence Review. – 2023. – Vol. 56. – P. 5545–5589. DOI: 10.1007/s10462-022-10283-5
7. Deep ensemble learning for accurate retinal vessel segmentation / [L. Du, H. Liu, L. Zhang et al.] // Computers in Biology and Medicine. – 2023. – Vol. 158. – P. 106829. DOI: 10.1016/j.compbiomed.2023.106829
8. Application of transfer learning and ensemble learning in image-level classification for breast histopathology / [Y. Zheng, C. Li, X. Zhou et al.] // Intelligent Medicine. – 2023. – Vol. 3, No. 2. – P. 115–128. DOI: 10.1016/j.imed.2022.05.004
9. Ensemble learning driven computer-aided diagnosis model for brain tumor classification on magnetic resonance imaging / [J. Vaiyapuri, S. Mahalingam, H. A. M. Ahmad et al.] // IEEE Access. – 2023. – Vol. 11. – P. 91398–91406. DOI: 10.1109/ACCESS.2023.3306961
10. Skin cancer detection using ensemble of machine learning and deep learning techniques / [J. V. Tembhurne, N. Hebbar, H. Y. Patil, et al.] // Multimedia Tools and Applications. – 2023. – Vol. 82. – P. 27501–27524. DOI: 10.1007/s11042-023-14697-3
11. Al-Zebari A. Ensemble convolutional neural networks and transformer-based segmentation methods for achieving accurate sclera segmentation in eye images / A. Al-Zebari // Signal, Image and Video Processing. – 2024. – Vol. 18. – P. 1879–1891. DOI: 10.1007/s11760-023-02891-7
12. Iqball T. Weighted ensemble model for image classification / T. Iqball, M. A. Wani // International Journal of Information Technology. – 2023. – Vol. 15. – P. 557–564. DOI: 10.1007/s41870-022-01149-8
13. Pyramid attention network for image restoration / [Y. Mei, Y. Fan, Y. Zhang et al.] // International Journal of Computer Vision. – 2023. – Vol. 131. – P. 3207–3225. DOI: 10.1007/s11263-023-01843-5
14. Multi-scale feature pyramid fusion network for medical image segmentation / [B. Zhang, Y. Wang, C. Ding et al.] // International Journal of Computer Assisted Radiology and Surgery. – 2023. – Vol. 18. – P. 353–365. DOI: 10.1007/s11548-022-02738-5
15. Hyperspectral image instance segmentation using spectral-spatial feature pyramid network / [L. Fang, Y. Jiang, Y. Yan et al.] // IEEE Transactions on Geoscience and Remote Sensing. – 2023. – Vol. 61. – P. 1–13. DOI: 10.1109/TGRS.2023.3240481
16. CANet: Context aware network with dual-stream pyramid for medical image segmentation / [X. Xie, W. Zhang, X. Pan et al.] // Biomedical Signal Processing and Control. – 2023. – Vol. 81. – P. 104437. DOI: 10.1016/j.bspc.2022.104437
17. Large-scale tissue histopathology image segmentation based on feature pyramid / [P. Qin, J. Chen, J. Zeng et al.] // EURASIP Journal on Image and Video Processing. – 2018. DOI: 10.1186/s13640-018-0320-8
18. Automatic tuberculosis screening using chest radiographs / [S. Jaeger, A. Karargyris, S. Candemir et al.] // IEEE Transactions on Medical Imaging. – 2014. – Vol. 33, No. 2. – P. 233–245. DOI: 10.1109/TMI.2013.2284099
19. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration / [S. Candemir, S. Jaeger, K. Palaniappan et al.] // IEEE Transactions on Medical Imaging. – 2014. – Vol. 33, No. 2. – P. 577–590. DOI: 10.1109/TMI.2013.2290491


DOI 10.15588/1607-3274-2024-4-11

UDC 004.8

# AIRCRAFT DETECTION WITH DEEP NEURAL NETWORKS AND CONTOUR-BASED METHODS

**Radionov Y. D.** – Postgraduate student of Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

**Kashtan V. Yu.** – PhD, Associate Professor, Associate Professor of Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

**Hnatushenko V. V.** – Dr. Sc., Professor, Head of Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

**Kazymyrenko O.V.** – Postgraduate student of Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

## ABSTRACT

**Context.** Aircraft detection is an essential task in the military, as fast and accurate aircraft identification allows for timely response to potential threats, effective airspace control, and national security. The use of deep neural networks improves the accuracy of aircraft recognition, which is essential for modern defense and airspace monitoring needs.

**Objective.** The work aims to improve the accuracy of aircraft recognition in high-resolution optical satellite imagery by using deep neural networks and a method of sequential boundary traversal to detect object contours.

**Method.** A method for improving the accuracy of aircraft detection on high-resolution satellite images is proposed. The first stage involves collecting data from the HRPlanesv2 dataset containing high-precision satellite images with aircraft annotations. The second stage consists of preprocessing the images using a sequential boundary detection method to detect object contours. In the third stage, training data is created by integrating the obtained contours with the original HRPlanesv2 images. In the fourth stage, the YOLOv8m object detection model is trained separately on the original HRPlanesv2 dataset and the dataset with the applied preprocessing, which allows the evaluation of the impact of additional processed features on the model performance.

**Results.** Software that implements the proposed method was developed. Testing was conducted on the primary data before preprocessing and the data after its application. The results confirmed the superiority of the proposed method over classical approaches, providing higher aircraft recognition accuracy. The mAP50 index reached 0.994, and the mAP50-95 index reached 0.864, 1% and 4.8% higher than the standard approach.

**Conclusions.** The experiments confirm the effectiveness of the proposed method of aircraft detection using deep neural networks and the process of sequential boundary traversal to detect object contours. The results indicate this approach's high accuracy and efficiency, which allows us to recommend it for use in research related to aircraft recognition in high-resolution images. Further research could focus on improving image preprocessing methods and developing object recognition technologies in machine learning.

**KEYWORDS:** machine learning, image and contour recognition, optical image preprocessing, high-resolution imagery, aircraft detection.

## ABBREVIATIONS

AI is an Artificial Intelligence;
CNN is a Convolutional Neural Network;
FP is a False Positive;
FN is a False Negative;
FPS is a Frames Per Second;
IoU is an Intersection Over Union;
ML is Machine Learning;
SAR is a Synthetic-Aperture Radar;
TN is a True Negative;
TP is True Positive;
UAV is Unmanned Aerial Vehicle.

## NOMENCLATURE

*AP* is an average precision for a single class;
*mAP* Is a mean average precision;
*Precision* is a is the fraction of TP detections among all detections made at a particular IoU threshold;
*Recall* is the fraction of TP detections found among all possible detections made at a particular threshold.

## INTRODUCTION

In the contemporary information landscape, characterized by an exponential growth in data, object recognition tasks have gained immense prominence. Artificial intelligence and machine learning have emerged as pivotal technologies for addressing these challenges. Beyond their widespread application in socioeconomic spheres, these technologies hold immense significance in the military domain. The proliferation of satellite imagery, coupled with the surge in the utilization of unmanned aerial vehicles (UAVs) of varying scales in military operations, has generated a deluge of visual information far exceeding the processing capabilities of human analysts.

Object recognition in the information age is a critical task with far-reaching implications for both civilian and military applications. AI and ML technologies offer a promising approach to address the challenges posed by the massive volume, complexity, and variability of visual data.

**The object of study** is aircraft detection on high-resolution satellite images using machine learning methods.


DOI 10.15588/1607-3274-2024-4-12

**The subject of study** is the method for aircraft detection on high-resolution satellite images using machine learning methods.

**The purpose of the work** is to increase the accuracy of aircraft detection on high-resolution satellite images using machine learning methods.

## 1 PROBLEM STATEMENT

The burgeoning availability of high-resolution satellite imagery has opened a new frontier for automated aircraft detection in diverse applications, from surveillance to environmental monitoring. However, achieving high recognition accuracy remains a formidable challenge. This challenge stems from the inherent complexities of satellite imagery, including variable backgrounds, fluctuating illumination conditions, and indistinct object edges in raw images. These factors significantly impede the feature extraction capabilities of deep learning models, often resulting in the misidentification of aircraft. Consequently, a critical need exists for novel methodologies that enhance the effectiveness of deep learning models for aircraft detection in high-resolution satellite imagery. This research addresses this gap by proposing a novel preprocessing approach that leverages the power of contour detection techniques. This method aims to improve the performance of deep learning models in aircraft recognition tasks by strategically accentuating the boundaries of aircraft objects and suppressing background clutter.

## 2 REVIEW OF THE LITERATURE

Deep learning techniques, especially convolutional neural networks (CNNs), have proven highly effective in object recognition and classification tasks in recent years. Researchers actively apply these approaches to analyze aerospace imagery obtained from uncrewed aerial vehicles, satellites, and other remote sensing platforms. These tasks have gained significant importance in a wide range of applications, including object detection in autonomous vehicles [1], facial recognition [2], medical image analysis [3], and remote sensing [4].

Machine learning techniques have notably advanced remote sensing applications, particularly in automatically extracting water bodies from satellite imagery. In work [5], machine learning methods for water body detection using Sentinel-2 imagery were investigated, showing significant improvements in the accuracy and efficiency of remote sensing techniques. Specifically, the paper [6] proposed an enhancement to the YOLOv5 architecture, resulting in a 3.5% improvement in model accuracy.

Also, in another paper [7], authors proposed a framework that tackles small object detection in high-resolution remote sensing images. It utilizes a deconvolutional module to refine feature maps and recover spatial information lost during pooling layers. Additionally, a squeeze-and-excitation attention mechanism focuses on informative features crucial for small object detection. This combination achieves high accuracy in detecting small aircraft within complex backgrounds.

RefineContourNet is a ResNet-based multi-path refinement CNN specifically designed for object contour detection [8] could be explored for refining aircraft object boundaries after initial detection by deep learning models.

In another study [9], authors developed an algorithmic technique for aircraft segmentation based on fuzzy logic that achieves an accuracy of 92.5%.

YOLO-extract algorithm [10] shows faster convergence speed and reduces the calculation amount by 45.3GFLOPs and a number of parameters by 10.526 million. At the same time, the algorithm increases mAP by 8.1% and detection speed by 3 times in aircraft detection compared to YOLOv5.

The YOLO-class model [11] represents an enhanced version of the YOLO-extract model, incorporating modifications to the network architecture. These changes significantly improved detection accuracy, increasing from 0.608 to 0.704, and in FPS, rising from 36.16 to 39.598, when compared to the original YOLO-extract.

YOLO architecture can be used and improved for aircraft detection using SAR data [12]. By adding Attention-Efficient Layer Aggregation Network-Head (A-ELAN-H) module that prioritizes essential features for improved accuracy authors increased *mAP50* by 2.1% and *mAP50-95* by 1.9%.

The current body of research has predominantly concentrated on one of two approaches: object detection or contour detection. While these methods have achieved notable success individually, their integration remains an underexplored area of research.

## 3 MATERIALS AND METHODS

The aircraft detection technology proposed in the paper consists of four stages, as shown in Fig. 2.

The first stage consists of downloading the HRPlanesv2 [13] dataset the second iteration of HRPlanes dataset. It contains 2120 very high-resolution Google Earth images, with a total of 14335 aircraft labeled. The dataset has been divided into three parts: 70% for training, 20% for validation, and 10% for testing.

The second stage is retrieving contours using OpenCV contour detection. A technique, based on the border following algorithm [14], for identifying and analyzing connected regions within an image that exhibit similar intensity or color characteristics. It analyzes the binary image to identify connected components, where all pixels share the same intensity (white in our case) and are adjacent to each other. These connected components represent the object boundaries or contours. These regions often correspond to distinct shapes or objects present in the image.

The four boundaries of an image are referred to as its frame. An image with width w and height h can be represented as a matrix of order $h \times w$, composed of individual pixels. The rows 1 and *h*, and the columns 1 and *w*, form the image frame. A pixel with a gray value of zero is defined as a zero pixel, while a pixel with a gray value of one is termed a one pixel. In this algorithm, the frames of the binarized image are treated as zero pixels. If the frame


DOI 10.15588/1607-3274-2024-4-12

of the input image contains any one pixel, they are converted to zero pixels.

An example of Suzuki's algorithm is shown in Fig. 1, where pixels with the same absolute value are associated with the same boundary. The relationships between each boundary are illustrated on the right side of the figure. In this context, *ob* refers to the outer border, *hb* denotes the hole border, and the parent border indicates that the outer layer acts as the parent of the inner layer.

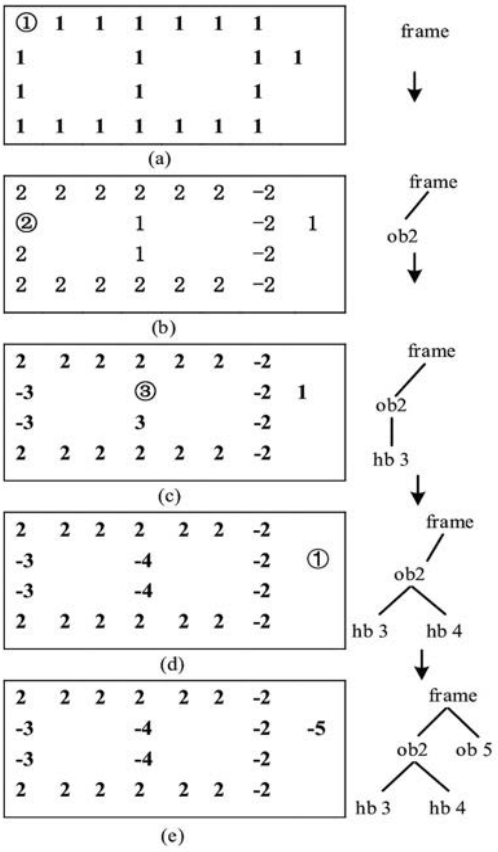



Figure 1 – Example diagram of border-following algorithm [15]

The circled pixel points in (a–e) correspond to the border descriptions for each link on the right.

The third stage is data pre-processing by combining contours with the original image. During the fourth stage, we opted to train the YOLOv8m object detection model. The architecture of the model is shown in Fig. 3.

Model selection was motivated by its favorable balance between model complexity (295 layers and 25.9 million parameters) and reported performance. This characteristic allowed us to balance object detection accuracy and computational efficiency.

The YOLOv8 architecture adheres to a modular design principle and can be divided into two primary components: the backbone and the head.

The backbone network serves as the foundation for feature extraction. It consists of 53 convolutional layers enhanced with cross-stage partial connections. YOLOv8 offers flexibility by employing a variety of backbone options, including CSPDarknet53 and EfficientDet. This selection allows for a trade-off between the capability to extract informative features and the associated computational complexity.

The head generates predictions based on the features extracted by the backbone network and the neck architecture. It predicts bounding box coordinates, objectness scores, and class probabilities for each anchor box associated with a grid cell. The architecture uses anchor boxes to predict objects of different shapes and sizes efficiently.

## 4 EXPERIMENTS

In the paper, we tested the proposed method on the HRPlanesv2 dataset and compared different contour detection techniques:

1. Canny edge detection. A multi-stage algorithm designed for robust edge extraction in images. It achieves a balance between high edge detection rates, accurate localization of edges, and minimal false positives. The algorithm employs a series of steps, including noise reduction through Gaussian filtering, calculation of image gradients, non-maximum suppression for thinning edges, and hysteresis thresholding for robust edge selection.

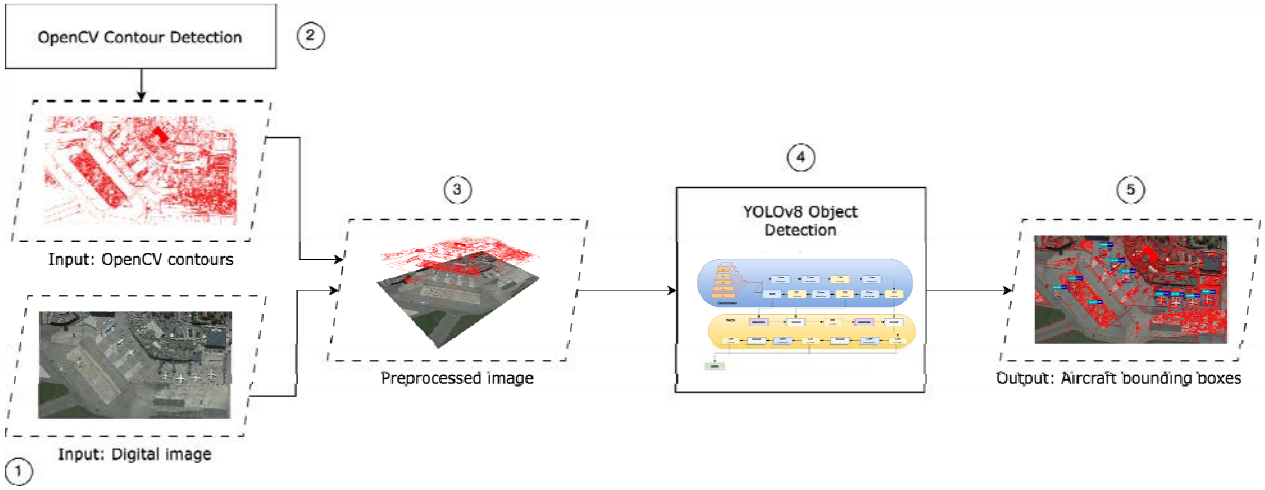



Figure 2 – Diagram of the proposed method


DOI 10.15588/1607-3274-2024-4-12

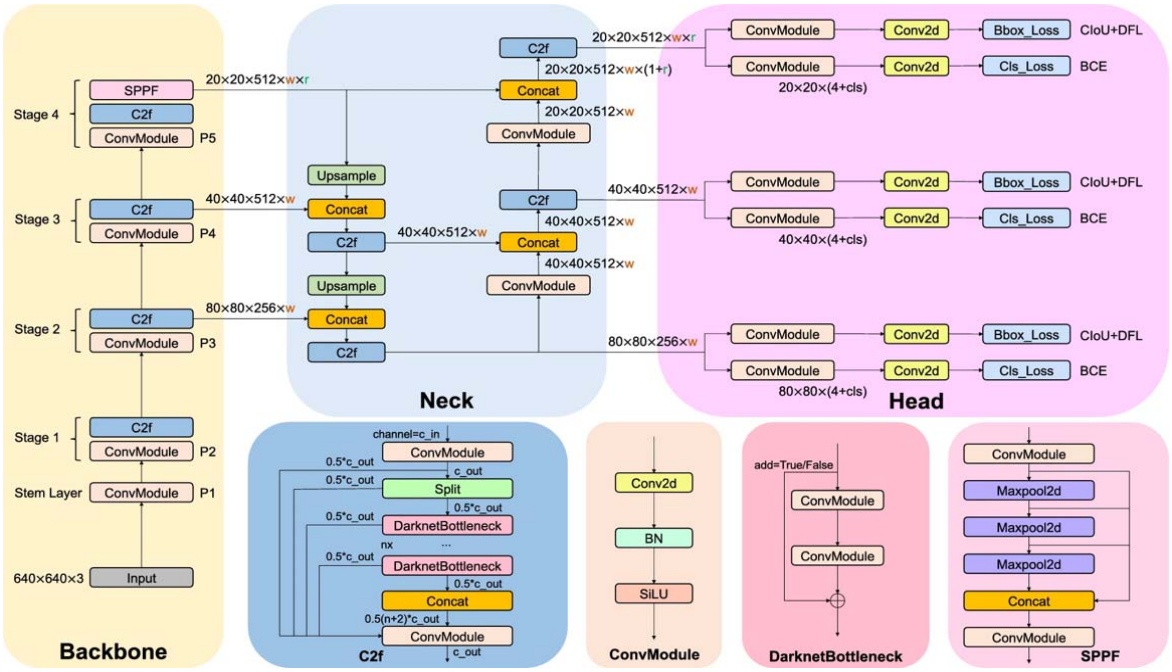



Figure 3 – YOLOv8 architecture [16]

2. Border following algorithm for contour detection.

3. The Laplacian. A mathematical operator is employed in image processing for edge detection and image sharpening. It calculates the second derivative of the image intensity, highlighting regions with rapid intensity changes, which frequently correspond to edges. Positive Laplacian values signify regions where intensity brightens towards the center (convex regions), while negative values indicate regions where intensity darkens (concave regions). A zero value indicates flat areas or edges where the direction of intensity changes flips.

4. The Sobel filter. A fundamental tool used for edge detection in image processing. It approximates the image gradient, representing the direction and magnitude of intensity change within the image. The filter utilizes two small kernels, one designed to detect horizontal edges and another for vertical edges, to estimate these gradients. The Sobel filter plays a crucial role in edge detection and feature extraction and serves as a building block for the Canny edge detection algorithm.

5. Canny and Laplacian operators applied sequentially.

Since YOLO measures the model's accuracy in terms of its ability to identify and locate objects of interest in an image accurately, the *mAP* was used. *mAP* is a combination of precision and recall values calculated over multiple confidence thresholds, also called the Intersection over Union threshold. Varying the IoU threshold will result in different True Positives and False Positives.

Precision is the fraction of TP detections among all detections made at a particular IoU threshold by the formula (1):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

Recall by the formula (2) is the fraction of TP detections found among all possible detections made at a particular threshold:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

The general formula for calculating *mAP* is shown in formula (3):

$$\text{mAP} = \frac{1}{\text{N}} \sum_{n=1}^{N} AP_i, \tag{3}$$

where N is the number of object classes, and $AP_i$ is the average precision for the *i*-th class. It's common to denote what IoU thresholds are used with digits following the *mAP*, e.g. *mAP50-95* uses a range of IoU thresholds from 0.5 to 0.95 with a 0.05 step size while *mAP95* uses a IoU threshold of 0.95.

The average precision for a given class is determined by first sorting all detected objects in descending order based on their confidence scores, then calculating the precision and recall at each threshold, and finally finding the area under the precision-recall curve by integrating it across all possible thresholds, as outlined in formula (4):

$$\text{Average Precision (AP)} = \int_{r=0}^{1} p(r)\text{d}r \tag{4}$$

In machine learning and computer vision, a confusion matrix is an essential tool for evaluating the performance of classification models. This matrix provides a clear


DOI 10.15588/1607-3274-2024-4-12

visualization of classification results, helping to assess the model's accuracy. It records the number of correct and incorrect classifications for each category. Specifically, TP refers to correctly classified positive instances, while TN denotes correctly classified negative cases. FP represents instances incorrectly classified as positive, and FN refers to instances mistakenly classified as negative.

## 5 RESULTS

Table 1 summarizes the accuracy metrics the object detection model achieved on datasets that underwent various pre-processing techniques. The presented metrics provide insights into the effectiveness of different pre-processing methods in enhancing the model's ability to detect objects within the datasets.

The original, unprocessed dataset achieved the highest *Recall* value and the second-highest *mAP50-95*. This suggests that the raw images already contain sufficient information for the model to identify a significant portion of aircraft present.

Its superior performance in Precision, *mAP50*, and *mAP50-95* demonstrates this. The results suggest that the extracted contours significantly improve the model's ability to accurately detect and distinguish aircraft from background elements. Consequently, this technique was chosen for the proposed method.

The Laplacian operator and Sobel filter pre-processing techniques demonstrated descent performance, but significantly lower than the original dataset and border-following contour detection. However, these methods led to a considerable loss of information compared to the original images. This indicates that, although they may emphasize potential edges, they could introduce artifacts or eliminate essential details needed for accurate object detection.

The Canny edge detection algorithm produced the lowest performance across all metrics. This could be attributed to the presence of noise or inconsistencies in the high-resolution satellite imagery.

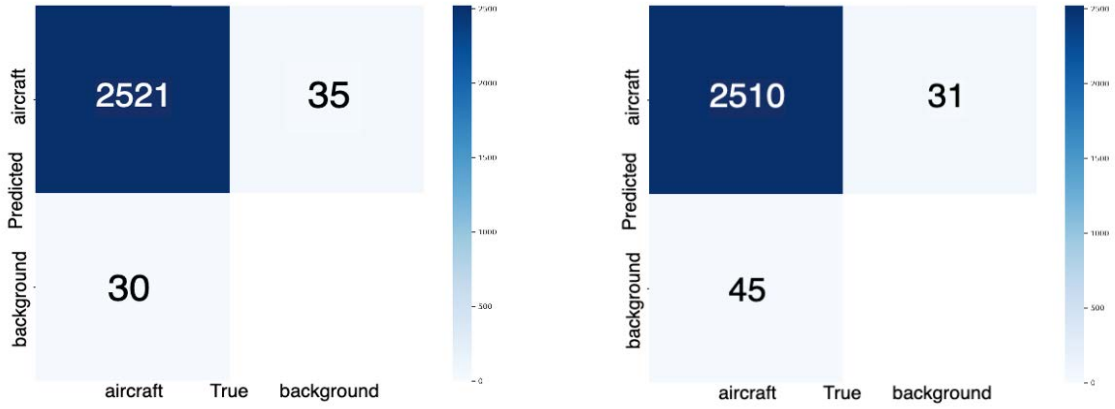
The combined Canny-Laplacian approach yielded average results, falling between the single Canny and Laplacian pre-processing performance. While it outperformed Canny alone, it remained less effective than Laplacian pre-processing. This suggests that the combined approach might not have effectively leveraged the strengths of both techniques, potentially introducing redundant or conflicting information for the model.

These findings highlight the importance of selecting appropriate pre-processing techniques for specific image datasets and tasks. While edge detection can enhance feature extraction, the choice of method needs to balance edge delineation with information preservation to achieve optimal performance in deep learning models for object recognition.

The confusion matrix, presented in Fig. 4, illustrates the performance of the classification models by displaying the number of correctly and incorrectly classified objects across each category.

Table 1 – Aircraft detection accuracy metrics

| Pre-processing type | Precision | Recall | mAP50 | mAP50-95 |
|---|---|---|---|---|
| None | 0.990522 | **0.981679** | 0.991079 | 0.831246 |
| Canny | 0.902949 | 0.782908 | 0.859798 | 0.592251 |
| Border-following | **0.996403** | 0.975793 | **0.994272** | **0.863925** |
| Laplacian | 0.951500 | 0.885149 | 0.958168 | 0.713322 |
| Sobel | 0.965968 | 0.958904 | 0.982615 | 0.809901 |
| Canny + Laplacian | 0.948966 | 0.827006 | 0.897425 | 0.654177 |




a b

Figure 4 – Confusion matrix for: a – original dataset; b – a dataset with the proposed technology

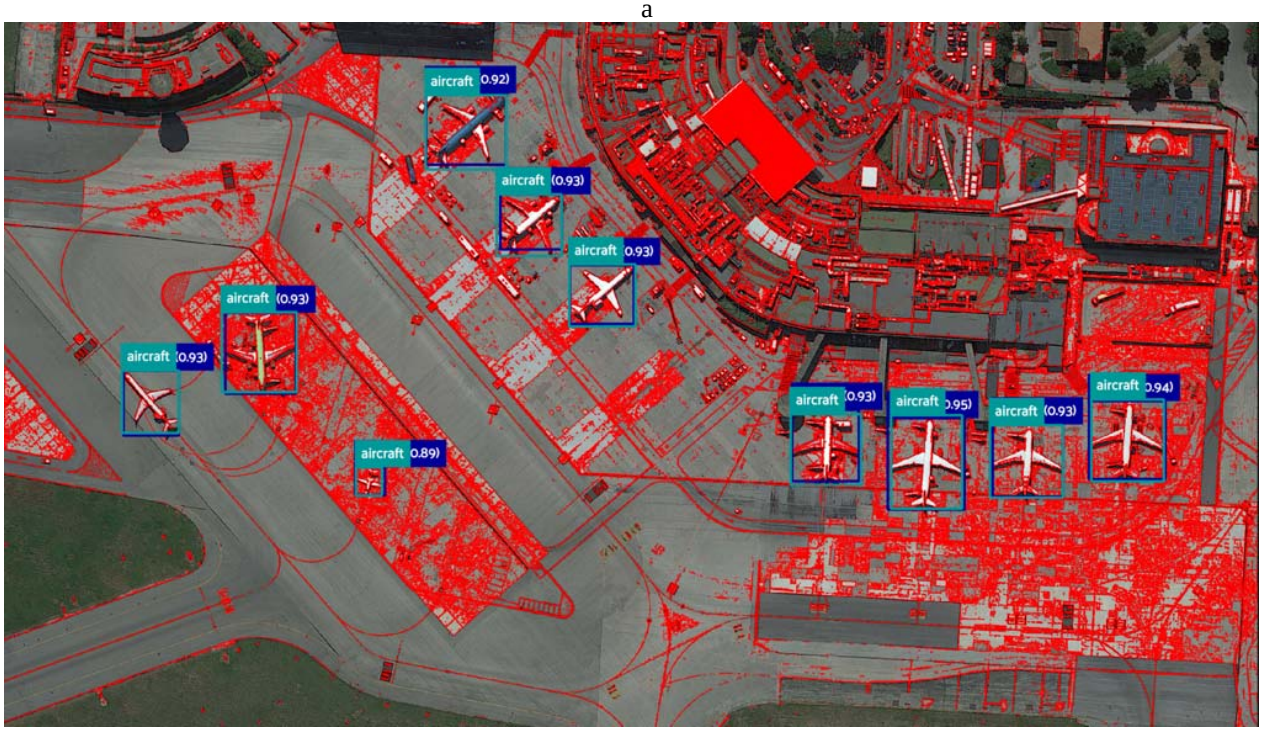
The results shown in Fig. 4 demonstrate the comparison of confusion matrices for two approaches: using the original dataset (Fig. 4a) and the modified dataset using the proposed technology (Fig. 4b). In the original dataset, the number of correctly classified objects in the "airplane" class was 2,521, with 71 false positives and 34 false nega-


DOI 10.15588/1607-3274-2024-4-12

tives. After implementing the proposed method, the number of false positives dropped to 31, reflecting an improvement in accuracy. However, the false negatives increased to 45, which means a specific decrease in completeness. Thus, the proposed technology improves the accuracy of classifying objects of the “airplane” class, but this is achieved at the expense of a slight decrease in the completeness index.

Fig. 5 illustrates the detection outcomes for the original dataset without pre-processing (a) and the dataset after applying the border-following contour detection method (b). In both images, cyan bounding boxes indicate ground truth annotations, while blue bounding boxes show the detections made by the respective model.


a




b

Figure 5 – Results of proposed technology: a – original dataset [14]; b – dataset with proposed technology


DOI 10.15588/1607-3274-2024-4-12

## 6 DISCUSSION

The proposed aircraft detection technology, leveraging edge detection pre-processing in a four-stage pipeline (Fig. 2), improved recognition accuracy compared to a baseline model trained on raw images. By incorporating pre-processed images with accentuated aircraft boundaries, the model is better equipped to distinguish aircraft from background clutter, improving recognition performance.

Our investigation into various edge detection techniques yielded valuable insights. The border-following algorithm, implemented in OpenCV, emerged as the most effective. Intriguingly, the combined Canny-Laplacian approach showed promise, suggesting potential for further optimization in future research to explore potential synergies between different techniques. While the proposed method demonstrates a clear advancement in aircraft detection using high-resolution satellite imagery, there are limitations to consider. The evaluation was conducted using a single dataset (HRPlanesV2). Future work should expand upon these findings by incorporating a broader range of datasets encompassing diverse image characteristics and backgrounds. This will provide a more comprehensive understanding of the generalizability and robustness of the proposed approach across different scenarios. Additionally, exploring alternative deep learning models with architectures specifically designed for high-resolution image recognition tasks could potentially yield further improvements. Furthermore, investigating even more advanced edge detection techniques, leveraging recent advancements in deep learning for image processing, could be a fruitful avenue for future research.

In conclusion, this research effectively demonstrates the value of incorporating edge detection pre-processing into the aircraft detection pipeline. The proposed method achieved superior performance in recognizing aircraft within high-resolution satellite imagery. The findings pave the way for further exploration and refinement, ultimately contributing to developing increasingly accurate and robust automated aircraft detection systems.

## CONCLUSIONS

This study investigated the impact of pre-processing techniques on the performance of a deep-learning model for aircraft detection in high-resolution imagery.

Our results show that pre-processing techniques can significantly improve a model's precision in object detection tasks. The findings emphasize how pre-processing enhances the model's ability to extract critical features from the data, leading to greater detection accuracy. Additionally, the study highlights the importance of understanding the interaction between pre-processing methods and the selected deep-learning model architecture.

**The scientific novelty** o of the obtained results is that, for the first time, a method is proposed to improve the accuracy of aircraft recognition on high-resolution satellite images, combining deep neural networks with the technique of sequential boundary traversal to detect object contours. The developed approach allows automating the process of aircraft detection and provides high recognition accuracy, which is a significant improvement over the existing classical method. The implementation of this method allows for efficient image processing, increasing the accuracy of object detection by optimizing the use of preprocessing and modern machine-learning algorithms.

**The practical significance** lies in developing software that implements the proposed method. The experiments have confirmed the effectiveness of the new approach for recognizing aircraft on satellite images. The research results suggest using this method for practical airspace monitoring and aviation security tasks. Further improvement of image preprocessing methods and object recognition technologies can increase the efficiency of these systems in various practical applications.

**Prospects for further research** are to improve image preprocessing methods. In addition, it is advisable to explore the possibilities of expanding the functionality of the developed software by integrating new neural network architectures and advanced machine learning algorithms to improve the accuracy and speed of real-time object recognition. It also promises to apply the proposed approach to other objects and images, which can expand its practical application and increase its versatility in various fields.

## REFERENCES

1. Tarmizi I. A., Aziz A. A.Vehicle Detection Using Convolutional Neural Network for Autonomous Vehicles, *International Conference on Intelligent and Advanced System, ICIAS 2018.* Kuala Lumpur, Malaysia, 2018, pp. 1–5. – DOI: 10.1109/ICIAS.2018.8540563.
2. Li L., Mu X., Li S., Peng H. A Review of Face Recognition Technology, *IEEE Access*, 2020, Vol. 8, pp. 139110–139120. DOI: 10.1109/ACCESS.2020.3011028.
3. Galić I., Habijan M., Leventić H., Romić K. Machine Learning Empowering Personalized Medicine: A Comprehensive Review of Medical Image Analysis Methods, *Electronics*, 2023, Vol. 12, № 21, P. 4411. DOI: 10.3390/electronics12214411.
4. Hnatushenko V., Kashtan V. Automated pansharpening information technology of satellite images, *Radio Electronics, Computer Science, Control*, 2021, № 2, pp. 123–132. DOI 10.15588/1607-3274-2021-2-13.
5. Kashtan V., Hnatushenko V. Machine learning for automatic extraction of water bodies using Sentinel-2 imagery, *Radio Electronics, Computer Science, Control*, 2024, № 1, pp. 118–127. DOI: 10.15588/1607-3274-2024-1-11
6. Zhou F., Deng H., Xu Q., Lan X. CNTR-YOLO: Improved YOLOv5 Based on ConvNext and Transformer for Aircraft Detection in Remote Sensing Images, *Electronics*, 2023, Vol. 12, № 12, P. 267. DOI: 10.3390/electronics12122671.
7. Zhou L., Yan H., Shan Y., Zheng Ch., Liu Y., Zuo X., Qiao B. Aircraft Detection for Remote Sensing Images Based on Deep Convolutional Neural Networks, *Journal of Electrical and Computer Engineering*, 2021, Vol. 2021, pp. 1–16. DOI: 10.1155/2021/4685644.
8. Kelm. A. P., Rao V. S., Zoelzer U. Object Contour and Edge Detection with RefineContourNet, *Computer Analysis of Images and Patterns. CAIP 2019, Salerno, 2–6 September, 2019*. Springer, Cham, 2019, pp. 246–258. DOI: 10.1007/978-3-030-29888-3_20.


DOI 10.15588/1607-3274-2024-4-12

9. Berezina S., Solonets O., Lee K., Bortsova M. An information technique for segmentation of military assets in conditions of uncertainty of initial data, *Information Processing Systems*, 2021, №4(167), pp. 6–18. DOI: 10.30748/soi.2021.167.01.
10. Liu Z., Gao Y., Du Q., Chen M., Lv W. YOLO-Extract: Improved YOLOv5 for Aircraft Object Detection in Remote Sensing Images, *IEEE Access*, 2023, Vol. 11, pp. 1742–1751. DOI: 0.1109/ACCESS.2023.3233964.
11. Liu Z., Gao Y., Du Q. YOLO-Class: Detection and Classification of Aircraft Targets in Satellite Remote Sensing Images Based on YOLO-Extract, *IEEE Access*, 2023, Vol. 11. pp. 109179–109188. DOI: 10.1109/ACCESS.2023.3321828.
12. Chen J., Shen Y., Liang Y., Wang Z., Zhang Q. YOLO-SAD: An Efficient SAR Aircraft Detection Network, *Applied Sciences*, 2023, Vol. 14, № 7, P. 3025. DOI: 10.3390/app14073025.
13. Unsal D. HRPlanesv2 – High Resolution Satellite Imagery for Aircraft Detection, *Zenodo*, 2022. DOI: 10.5281/ZENODO.7331974.
14. Suzuki S., Be K.Topological structural analysis of digitized binary images by border following, *Computer Vision, Graphics, and Image Processing*, 1985, Vol. 30, № 1, pp. 32–46. DOI: 10.1016/0734-189X(85)90016-7.
15. Song X., Zhang S., Yang J., Zhang J. Research on Luggage Package Extraction of X-ray Images Based on Edge Sensitive Multi-Channel Background Difference Algorithm, *Applied Sciences*, 2023, Vol. 13, № 21, P. 11981. DOI: 10.3390/app132111981.
16. Ju R-Y,. Cai W. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm, *Scientific Reports*, 2023, Vol. 13, №1, P. 20077. DOI: 10.1038/s41598-023-47460-7.

Received 02.07.2024.
Accepted 17.09.2024.

УДК 004.8

## РОЗПІЗНАВАННЯ ЛІТАКІВ ЗА ДОПОМОГОЮ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ ТА ВИЯВЛЕННЯ КОНТУРІВ

**Радіонов Є. Д.** – аспірант кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.
**Каштан В. Ю.** – канд. техн. наук, доцент, доцент кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.
**Гнатушенко В. В.** – д-р техн. наук, професор, завідувач кафедри інформаційних технологій та комп'ютерної інженерії, Національний технічний університет «Дніпровська політехніка», Дніпро, Україна.
**Казимиренко О. В.** – аспірант кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.

### АНОТАЦІЯ

**Актуальність.** Розпізнавання літаків є важливою задачею у військовій сфері, оскільки швидка та точна ідентифікація літальних апаратів дозволяє своєчасно реагувати на потенційні загрози, ефективно контролювати повітряний простір і підтримувати національну безпеку. Використання глибоких нейронних мереж підвищує точність розпізнавання літаків, що є важливим для сучасних потреб оборони та моніторингу повітряного простору.

**Мета роботи** – підвищення точності розпізнавання літаків на оптичних космічних знімках високої роздільної здатності за допомогою глибоких нейронних мереж та методу послідовного обходу меж для виявлення контурів об'єктів.

**Метод.** Запропоновано метод для підвищення точності розпізнавання літаків на супутникових знімках високої роздільної здатності. На першому етапі здійснюється збір даних із набору HRPlanesv2, що містить високоточні супутникові зображення з анотаціями літаків. Другий етап передбачає попередню обробку зображень за допомогою методу послідовного обходу меж для виявлення контурів об'єктів. На третьому етапі створюються навчальні дані шляхом інтеграції отриманих контурів з оригінальними зображеннями HRPlanesv2. На четвертому етапі модель виявлення об'єктів YOLOv8m тренується окремо на оригінальному наборі даних HRPlanesv2 та на наборі даних із застосованою попередньою обробкою, що дозволяє оцінити вплив додаткових оброблених характеристик на продуктивність моделі.

**Результати.** Розроблено програмне забезпечення, яке реалізує запропонований метод. Тестування проводилося як на первинних даних до попередньої обробки, так і на даних після її застосування. Результати підтвердили перевагу запропонованого методу над класичними підходами, забезпечуючи вищу точність розпізнавання літаків. Показник mAP50 досяг 0.994, а mAP50-95 – 0.864, що на 1% і 4,8% відповідно, вище, ніж у стандартного підходу.

**Висновки.** Проведені експерименти підтверджують ефективність запропонованого методу розпізнавання літаків за допомогою глибоких нейронних мереж та методу послідовного обходу меж для виявлення контурів об'єктів. Результати вказують на високу точність і ефективність цього підходу, що дозволяє рекомендувати його для використання в задачах, пов'язаних із розпізнаванням літаків на зображеннях високої роздільної здатності. Подальші дослідження можуть зосередитися на вдосконаленні методів попередньої обробки зображень і розвитку технологій розпізнавання об'єктів у машинному навчанні.

**КЛЮЧОВІ СЛОВА:** машинне навчання, розпізнавання образів та контурів, попередня обробка оптичних зображень, знімки високої роздільної здатності, розпізнавання літаків.

### ЛІТЕРАТУРА

1. Tarmizi I. A. Vehicle Detection Using Convolutional Neural Network for Autonomous Vehicles / I. A. Tarmizi, A. A. Aziz // International Conference on Intelligent and Advanced System, ICIAS 2018, Kuala Lumpur, Malaysia. – 2018. – P. 1–5. – DOI: 10.1109/ICIAS.2018.8540563.
2. A Review of Face Recognition Technology / [L. Li, X. Mu, S. Li, H. Peng] // IEEE Access. – 2020. – Vol. 8. – P. 139110–139120. DOI: 10.1109/ACCESS.2020.3011028.


DOI 10.15588/1607-3274-2024-4-12

3. Machine Learning Empowering Personalized Medicine: A Comprehensive Review of Medical Image Analysis Methods / [I. Galić, M. Habijan, H. Leventić, K. Romić] // Electronics. – 2023. – Vol. 12 № 21. – P. 4411. DOI: 10.3390/electronics12214411.
4. Hnatushenko V. Automated pansharpening information technology of satellite images / V. Hnatushenko, V. Kashtan // Radio Electronics, Computer Science, Control. – 2021. – № 2. – P. 123–132. DOI:10.15588/1607-3274-2021-2-13.
5. Kashtan V. Machine learning for automatic extraction of water bodies using Sentinel-2 imagery / V. Kashtan, V. Hnatushenko // Radio Electronics, Computer Science, Control. – 2024. – № 1. – P. 118–127. DOI: 10.15588/1607-3274-2024-1-11
6. CNTR-YOLO: Improved YOLOv5 Based on ConvNext and Transformer for Aircraft Detection in Remote Sensing Images / [F. Zhou, H. Deng, Q. Xu, X. Lan] // Electronics. – 2023. – Vol. 12, № 12. – P. 267. DOI: 10.3390/electronics12122671.
7. Aircraft Detection for Remote Sensing Images Based on Deep Convolutional Neural Networks / [L. Zhou, H. Yan, Y. Shan, Ch. Zheng et al.] // Journal of Electrical and Computer Engineering. – 2021. – Vol. 2021. – P. 1–16. DOI: 10.1155/2021/4685644.
8. Kelm. A. P. Object Contour and Edge Detection with RefineContourNet / A. P. Kelm, V. S. Rao, U. Zoelzer // Computer Analysis of Images and Patterns. CAIP 2019, Salerno, 2–6 September, 2019. – Springer, Cham, 2019. – P. 246–258. DOI: 10.1007/978-3-030-29888-3_20.
9. An information technique for segmentation of military assets in conditions of uncertainty of initial data / [S. Berezina, O. Solonets, K. Lee, M. Bortsova] // Information Processing Systems. – 2021. – №4(167). – P. 6–18. DOI: 10.30748/soi.2021.167.01.
10. YOLO-Extract: Improved YOLOv5 for Aircraft Object Detection in Remote Sensing Images / [Z. Liu, Y. Gao, Q. Du et al.] // IEEE Access. – 2023. – Vol. 11. – P. 1742–1751. DOI: 0.1109/ACCESS.2023.3233964.
11. Liu Z. YOLO-Class: Detection and Classification of Aircraft Targets in Satellite Remote Sensing Images Based on YOLO-Extract / Z. Liu, Y. Gao, Q. Du // IEEE Access. – 2023. – Vol. 11. – P. 109179–109188. DOI: 10.1109/ACCESS.2023.3321828.
12. Chen J. YOLO-SAD: An Efficient SAR Aircraft Detection Network / [J. Chen, Y. Shen, Y. Liang et al.] // Applied Sciences. – 2023. – Vol. 14, № 7. – P. 3025. DOI: 10.3390/app14073025.
13. Unsal D. HRPlanesv2 – High Resolution Satellite Imagery for Aircraft Detection / D. Unsal // Zenodo. – 2022. DOI: 10.5281/ZENODO.7331974.
14. Suzuki S. Topological structural analysis of digitized binary images by border following / S. Suzuki and K. Be // Computer Vision, Graphics, and Image Processing. – 1985. – Vol. 30, № 1. – P. 32–46. DOI: 10.1016/0734-189X(85)90016-7.
15. Research on Luggage Package Extraction of X-ray Images Based on Edge Sensitive Multi-Channel Background Difference Algorithm. / [X. Song, S. Zhang, J. Yang, J. Zhang] // Applied Sciences. – 2023. – Vol. 13, № 21. – P. 11981. DOI: 10.3390/app132111981.
16. Ju R-Y. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm / R.-Y. Ju, W. Cai // Scientific Reports. – 2023. – Vol. 13, № 1. – P. 20077. DOI: 10.1038/s41598-023-47460-7.


DOI 10.15588/1607-3274-2024-4-12

UDC 681.325

# POST PROCESSING OF PREDICTIONS TO IMPROVE THE QUALITY OF RECOGNITION OF WATER SURFACE OBJECTS

**Smolij V. M.** – Dr. Sc., Professor, Professor of the Department of Information systems and technologies, National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine.

**Smolij N. V.** –Student of the Department of Information systems and technologies, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.

**Mokriiev M. V.** – PhD, Associate professor of the Department of Information systems and technologies, National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine.

## ABSTRACT

**Context.** The significance of this work stems from the growing need for UAV technologies integrated with artificial intelligence, aimed at detecting and identifying objects on the surface of water bodies. Modern needs in water body monitoring, especially in the context of environmental monitoring, protection and resource management, require accurate and reliable solutions. This work demonstrates methods for improving the performance of neural networks and offers approaches for processing NN predictions, even if they are trained on irrelevant data, which increases the versatility and efficiency of the technology.

**Objective.** The goal of the work is to solve the problem of false recognition of objects on the surface of water bodies, which is due to a decrease in the accuracy threshold for the neural network. This provides more accurate and reliable detection, reducing the number of false positive predictions and increasing the efficiency of the system in general.

**Method.** It is proposed to add a stage of post-processing of NN predictions, which inherits concepts of min-max suppression used by YOLO models. This algorithm suppresses the re-detection of the object by the network and relies on the cross-sectional area of the detected rectangles. It uses a threshold value of 0.8 for the two points of the rectangle, which can effectively reduce the number of re-predictions and improve the accuracy.

**Results.** As a result of the implementation of the proposed algorithm and the script created on its basis, a result was achieved in which groups from several predictions are combined and filtered. The received data is stored in the database as found and detected objects. The proposed post-processing algorithm effectively removes redundant predictions while maintaining forecast accuracy. This ensures the reliability of the system and increases its performance in real conditions.

**Conclusions.** Detected images of objects on the surface of water bodies are stored in the database in the form of records with unique file name identifiers. After tests with pre-taken images algorithm proved it`s persistence against data duplication scenarios. This increases the efficiency and reliability of the monitoring system, ensuring accurate and timely detection of objects on the surface of water bodies.

**KEYWORDS:** UAV, detection, recognized objects, water surface, neural network, dataset, model, image distribution, error matrix, training metrics, magnification, image mosaicking, image post-processing, implementation script, mission log, database.

## ABBREVIATIONS

AI stands for artificial Intelligence;
FC is a flight controller;
UAV is a unmanned aerial vehicle;
YOLO is the You Only Look Once;
NN is a neural network;
RS is a remote sensing;
SPI is a serial peripheral interface;
IS is a information system;
ConvN is a deep convolution neural network;
DB is a database;
AP is a indicator of average accuracy – average accuracy;
IoU is a intersection over union;
NCU is a nonlinear model predictive control.

## NOMENCLATURE

*x*i, *y*i is the coordinates of the corresponding corner of the rectangle;
*TP* is a true positive;
*TN* is a true negative;
*FP* is a false positive;
*FN* is a false negative;
*GT* is a ground truth;
*P* is a prediction;
*r* is a recall value for one sample;
$\tilde{r}$ is a average recall value.

## INTRODUCTION

The importance of the project arises from the current need for artificial intelligence integration into UAV technologies. This article shows bright example in context of localization and identification problems for objects on the surface of water bodies.

**The object of study** may be defined as algorithm description for task of quantity reduction of false recognition cases and further implementation into information system.

The process of forming and downloading the algorithm and the created script based on it, a result was obtained where groups from several predictions are combined and filtered. As a result, these findings are saved in the database in the form they were obtained.

**The subject of study** is the implementation of the proposed algorithm and the created script based on it, a result was obtained where groups from several predictions are combined and filtered. As a result, detected objects are saved in the database in the form they were obtained.


DOI 10.15588/1607-3274-2024-4-13

**The purpose of the work** is to solve the problem of false recognition of objects caused by manipulations with train and process parameters for neural network.

## 1 PROBLEM STATEMENT

The current state of development of artificial intelligence technologies significantly expands the range of tasks that can be solved by robotics tools [1], [2]. Most needed areas of automatization for area of UAVs involve such tasks as mission completion time reduction, more efficient usage of computational powers and reduction of time of human intervention into UAV control processes [3], [4], [5]. Patrolling water bodies, searching, identifying and classifying objects on their surface using traditional methods is a complex process with high requirements for time and computational powers that needs to be enhanced with implementation of various modern techniques such as artificial intelligence and mode [6], [7], [8].

In the process of performing the mission, the UAV finds a certain object, determines its coordinates, with the help of a neural network, prediction rectangles are generated, characterized by *AP*, *Pinterp*, which must be filtered. The same object can be recognized by the neural network as several different objects, which causes redundancy and duplication of data in the DB.

It is necessary to develop an algorithm (procedure, technique, method) of filtering (elimination of redundancy) generated rectangles in order to avoid duplication of predictions. It is also necessary to test the proposed algorithm, check the efficiency of the work, determine the cut-off threshold value at which the object will be detected, but there will be no duplication. This coefficient is defined as the ratio of the cross-sectional area to the areas of two rectangles. It is necessary to expand the box with which the intersection is fixed to the borders of the intersecting box. This achieves preservation of detection quality while eliminating duplication.

It was found that the coordinates of the upper right and lower left corners of the rectangle are considered for each object. A set of rectangles at two points ($x1$, $y1$, $x2$, $y2$) associated with objects that can intersect with each other and thus create duplicate objects. The specified rectangles are the result of neural network predictions. The number of specified rectangles is inversely proportional to the prediction accuracy threshold.

The job task is to eliminate duplication while maintaining the accuracy of recognition. To do this, it is necessary to filter this set of rectangles using the proposed threshold filtering algorithm with an overlap threshold of 0.8, getting rid of duplicate markings and not reducing the number of detected objects. It is imperative to investigate the operability and efficiency of the proposed post-processing algorithm.

## 2 REVIEW OF THE LITERATURE

For modern unmanned aerial vehicles, great variety of tasks such as: classification, observation, research, and classification are getting redirected to certain automated subsystems, which may enhance performance thanks to the additional computational modules involved [9], [10], [11]. Such type of an approach is appropriate and widespread due to the availability of flight controllers on the market [12], [13]. Available options suggest great variety of sockets for external devices for variety of purposes: exchange of GPS data, OSD frames and other types of messages defined via MavLink. Such data can be clustered into data sets for the variety of purposes: prioritized task queuing and data processing routines [14], [15]. Computer vision systems implement various machine learning algorithms and complex data processing techniques for image processing and localization tasks The complexity of the structure of a computer vision system may vary depending on the type of task and the stack of technologies used [16], [17].

As well as any other information system, computer vision system involves such functionality as data collection processing and storage. Additional functions may vary from performing computations for neural network up to data generalization, clustering and image segmentation with further additions in the form of decision-making process. The design of UAV for computer vision tasks includes the following structural elements: reconnaissance equipment, UAV frame, propeller group and flight controller. It should be noted that the range of use of UAVs in water and coastal zones is very large and responsible, since the functions of object detection, observation, identification, classification, etc. directly related to the security, protection and territorial integrity of Ukraine. The purpose of the research is to determine the accuracy of object recognition, to find out the reasons for possible unsatisfactory recognition quality, to form means and strategies for improving the quality of image recognition of objects. It should be noted that effective algorithms for the detection and classification of objects on the water surface have not yet been created, apart from artificial intelligence and neural network tools based on it.

The technical system, which includes UAVs and AI tools, will allow relatively fast and high-quality processing of data frames representing any type of surface below UAV, thereby providing efficient observation functions performance, response and elimination of emergency situations with minimal human involvement [18]. The main task of the work is to detect and recognize the object in the image. Mentioned area suggests great variety of artificial intelligence scientific works and solutions to study that differ in the structure of the neural networks used, the dataset used, and are accordingly characterized by the quality of object recognition. These works mostly rely on such ML libraries as pytorch, theano, tensorflow, chainer, ultralytics. The main goal of the work is to create a data set regarding the recognized object, which in turn requires the use of the Ultralytics API library, which provides a wide functionality for dataset gathering and marking and an interface for variety of new and aout of date convolution neural networks from YOLO family. Due to the support of dynamic typing, automatic garbage collection and the availability of a library API, the software in


DOI 10.15588/1607-3274-2024-4-13

the work uses Python programming language. Source [1] explains in details (www.overleaf.com).

## 3 MATERIALS AND METHODS

The task of patrolling water surfaces and detecting as well as identifying objects relies on interaction with a neural network, the detailed workings, architecture, and training of which are discussed in [20]. For this particular neural network, it is necessary to develop tools for generating an on-board flight log and producing reports on the detected objects. The Ultralytics API was selected for model interaction. Additionally, it is important to mention that the OpenCV library was utilized for image processing due to its ability to easily transition from handling test data from media files to real-time camera data.

Since the task conditions require a difficult environment of use, namely working near water bodies and short missions with the transfer of the payload, it was decided to use a quadcopter. The basis for the choice was: a smaller number of parts to be serviced, compared to other members of the family, and optimal indicators of stability, speed and maximum indicators of the weight that can be carried. The use of the scheme of distributed current controllers was chosen due to high wear resistance and economy. Due to the predicted high rates of wear of the parts, this scheme will not replace the board with all four current controllers, but replace only the controller that has failed.

To build a UAV, you need to have a flight controller or the ability to write all the necessary functionality. Since writing your own flight controller is a time-consuming task, which, in terms of time and effort, far exceeds the defined time limits of the diploma project, we consider FCs available on the market. Existing flight controllers use a variety of software.

Software leaders include software from ArduPilot, PX4 – open source projects with a large audience of users, and DJI – licensed software used on more expensive UAV models and designed for flexible use with different UAV models and configuration variations. It should also be noted that ArduPilot is the most ‘general’ project, since the range of tasks to be solved for this project ranges from the design of UAV rotor group control systems to the programming of autopilot systems. The PX4 is only a draft autopilot system.

The family of platforms that use the ArduPilot code as firmware is quite large, and usually the decision to use a certain controller model is determined only by the user’s budget. Thus, the decision was made to use the Pixhawk 2.4.8 flight controller. Among its advantages are: the possibility of connecting 2x GPS to combine data and increase positioning accuracy, a built-in telemetry filtering system, a large number of peripheral interfaces supporting I2C, UART, SPI protocols, and a large number of drivers for peripheral devices that are not can be ignored when considering the perspective of development/complication of the created information system during its life cycle.

Also, a separate advantage is the presence of the MavLink protocol, which allows you to organize the exchange of messages with peripheral devices, and thus connect the on-board computer, which will be able to use such data of the flight controller as: location assessment, positioning quality, battery charge level, weather conditions for making own decisions about the continuation of the mission.

Since the flight controllers from Pixhawk do not provide functionality for deploying a neural network and a database, and the presence of a network card is an exception, it was accordingly decided to install an on-board microcomputer on the UAV for use as a control, processing, storage and information transfer device. This approach will allow you to get a full-fledged operating system, access to the package manager and, accordingly, to a convenient setting of the environment, will facilitate the process of writing and testing software.

The GPIO module will allow you to organize a connection with your own module for water intake. The built-in software application RealVNC and the ability to set a static IP address will provide an opportunity to connect to the UAV through the mobile network in the absence of such computer components as a monitor and keyboard. Built-in USB ports provide a convenient connection with a flight controller for data exchange using the MavLink protocol.

The use of a neural network on RaspberryPI is a typical step in the tasks of a limited number of resources for building a robotic information system for analyzing the state of water bodies and recognizing objects with the help of AI technologies. This gives an absolute advantage to this platform due to the large volume of available materials on solving problems and setting up processes.

Since the RaspberryPi has an OS with a package manager capable of installing all the necessary functionality with the help of sudo apt install commands, and because of the low cost of the hardware and the prevalence of a free license for the software, it was decided to use it.

Anything related to the use of AI can be classified into one of the following categories: application use, application development, model development, or infrastructure design (Fig. 1).

Research is carried out in the area of application development, since it is necessary to introduce artificial intelligence methods into the user environment, and to train models, since the requirements for the use of any specific model have not been set.

Since the experience of working with the ultralytics API was gained during the training, it was decided to use it. This module provides opportunities for training a neural network and for its integration into the application. Also, during the training process, statistics are collected and the metrics of a certain trained model are automatically built. This simplifies the process of iterating between training stages, evaluating and changing training conditions.


DOI 10.15588/1607-3274-2024-4-13

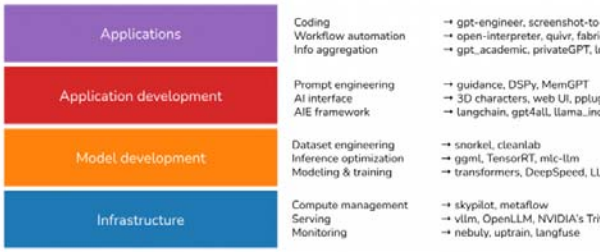



Figure 1 – AI technology stack


Figure 2 – Comparative characteristics of the accuracy of dimensional variations of two models depending on the type of task

The objects of observation are boats, ships, buoys, people and piles of garbage on water bodies. Accordingly, the task of the model is detection, since segmentation is not suitable for this type of work, although it can be used to improve accuracy or increase the functionality of the application being created.

According to the conditions of the task, in which there are significant limitations in computing power and the ability of UAVs to carry loads, it is necessary to use the second class of detection models, namely, single-stage detectors, which, accordingly, filter the image once. This approach reduces the accuracy of the forecast, reducing the amount of consumed computing resources. Also, in conditions when the dataset is collected from open sources, training anomalies are inevitable, which will lead to a decrease in accuracy even with a two-stage detector. Accordingly, the dual use of UAVs is expedient: both to perform the specified task and as a source of data for the formation of a better dataset.

Since it was necessary to achieve such a classification result that would not be completely wrong, with limited resources in the computer vision problem, it was decided to use the YOLOv8 model, as well as the API from Ultralytics, which was created in order to give the opportunity to train typical models on cash registers-volume datasets for various tasks.

A compilation of video footage from a fishing boat, video surveillance cameras of the pool, and photos from ecologists' boats was used as a dataset.

As it was mentioned in [24], 2 contenders were singled out among the family of YOLO models: YOLOv8, YOLOv5. The developers independently compared the accuracy of the classification, detection and segmentation tasks of the data models (Fig. 2).

Accordingly, it was decided to make our own tests for selected models in the Small dimension. Since we are facing an image detection task, the expected advantage towards YOLOv8 is 20% on 640x640 images.

For training, an example from the developers was used with arguments to improve the quality of training for the YOLOv8 and YOLOv5 models, respectively. Training was carried out for 100 epochs. The training results were evaluated based on such quality indicators as:

– the confusion matrix is a metric that characterizes how the neural network distinguishes between objects of different classes;

– P curve – a curve that characterizes the process of establishing the accuracy value as the number of correctly defined objects of the class in relation to the total number of defined objects of this class;

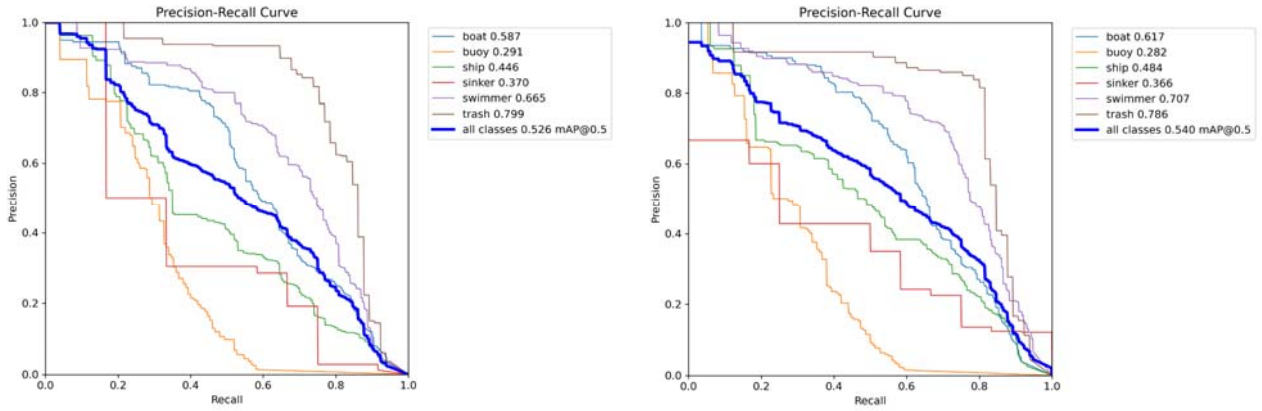
– R curve – a curve that characterizes the process of setting the recall-y value, which characterizes the ratio of the number of correctly defined objects of a certain class to their total number in the data set;

– mAP (:50, 50–95) curve – which characterizes the formation of the average accuracy value for the confidence threshold of 50% and in the interval from 50 to 95%, respectively. This metric provides an opportunity to evaluate how the model copes with different thresholds for accuracy/

Based on the results of training on the created dataset, the following diagrams describing the performance of 2 models were obtained. The results of the study of PR-metrics of the models, which characterize the change in the precision and recall of the parameters of the YOLOv5 and YOLOv8 models are presented in Fig. 3.

For boats, the degree of correlation between precision and recall indicators is 0.587 for the NN YOLOv5 model and 0.617 for the NN YOLOv8 model. The recognition of buoys occurs with a degree of correlation between the precision and recall indicators of 0.291 and 0.282 for YOLOv5 and YOLOv8 NN models, respectively. For ships, this indicator is 0.446 and 0.484, when recognizing drowning people – 0.370 and 0.366, when identifying swimmers, the correlation coefficient of precision and recall indicators is 0.665 and 0.707 (there is an advantage of the NN YOLOv8 model), and when recognizing on the other hand, with coefficients of 0.799 and 0.786, the YOLOv5 NN model is preferred in the bath of garbage images.


DOI 10.15588/1607-3274-2024-4-13




a b

Figure 3 – PR-metrics of models NN:
a – YOLOv5; b – YOLOv8

AP is the precision averaged over all completeness values between 0 and 1[25]. AP is interpreted as finding the area of the region below the precision-completeness curve. The mAP50-95 indicator is the average value of the average accuracy, calculated at different IoU threshold values, varying from 0.50 to 0.95. mAP@0.5 is the average accuracy calculated at an IoU threshold of 0.50. This is a measure of the accuracy of the model, taking into account only 'light' detections.

$$\mathrm{Pr}ecision = \frac{TP}{TP + FP}. \quad (1)$$

$$\mathrm{Re}call = \frac{TP}{TP + FN}. \quad (2)$$

$$IoU = \frac{area\ \ of\ \ overlap}{area\ \ of\ \ union}. \quad (3)$$

When studying the operation of the prediction filtering algorithm for each of the detected objects, the ratio of the planes of the rectangles shown in Fig. 4.


Figure 4 – Intersection of rectangles

$$AP = \int_0^1 p(r)dr. \quad (4)$$

Precision and recall values are normalized in range from 0 up to 1 that leads to AP being between values of 0 and 1 also. In additions it is a common practice to smooth zigzag pattern before calculating AP.

At every recall level, the precision value is adjusted to match the highest precision observed at any subsequent recall level on the graph:

$$p_{\mathrm{int}erp}(r) = \max_{\tilde{r} \geq r} p(\tilde{r}). \quad (5)$$

The curve spans all recall values ($r_1$, $r_2$, …), with a drop occurring each time the maximum precision value decreases. Such method provides exact area under pr-curve after pre-processing steps. With this change, we are measuring the exact area under the precision-recall curve after the zigzags are removed. General AP definition matches the definition of discrete integration of pr-curve.

There's no need for approximation or interpolation. We calculate AP by sampling $p(r_i)$ at each drop and summing the resulting rectangular areas:

$$AP = \sum (r_{n+1} - r_n) \cdot p_{\mathrm{int}erp}(r_{n+1}), \quad (6)$$

$$p_{\mathrm{int}erp}(r_{n+1}) = \max_{\tilde{r} \geq r_{n+1}} p(\tilde{r}). \quad (7)$$

The generalized index of precision and recall for all varieties of studied object classes is higher for YOLOv8 compared to YOLOv5 and is 0.540 versus 0.526, respectively.

The results of the study of RC-metrics of the models, which characterize the change in the recall and confidence of the parameters of the YOLOv5 and YOLOv8 models are presented in Fig. 5.

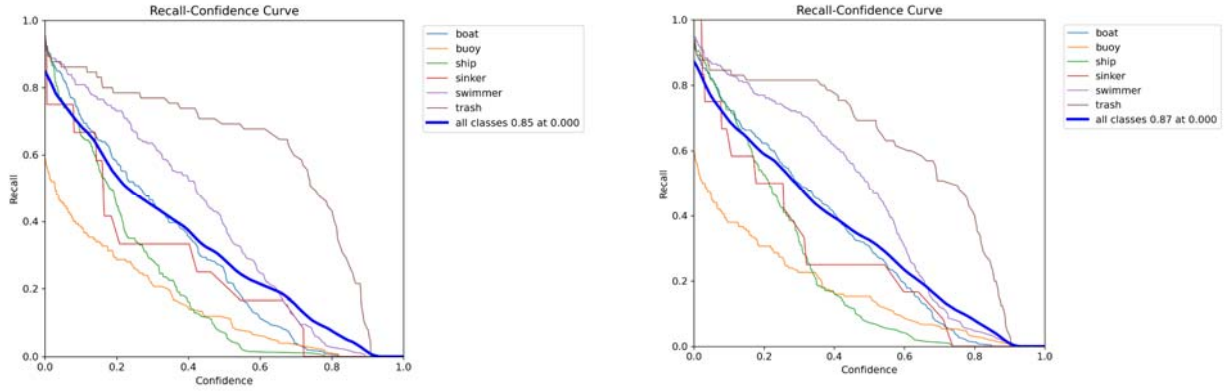
The generalized indices of recall and confidence for all varieties of the studied object classes have roughly equivalent values for YOLOv8 compared to YOLOv5 and are 0.87 versus 0.85, respectively.

The results of the study of PC-metrics of the models, which characterize the change in the precision and confidence of the parameters of the YOLOv5 and YOLOv8 models are presented in Fig. 6.


DOI 10.15588/1607-3274-2024-4-13

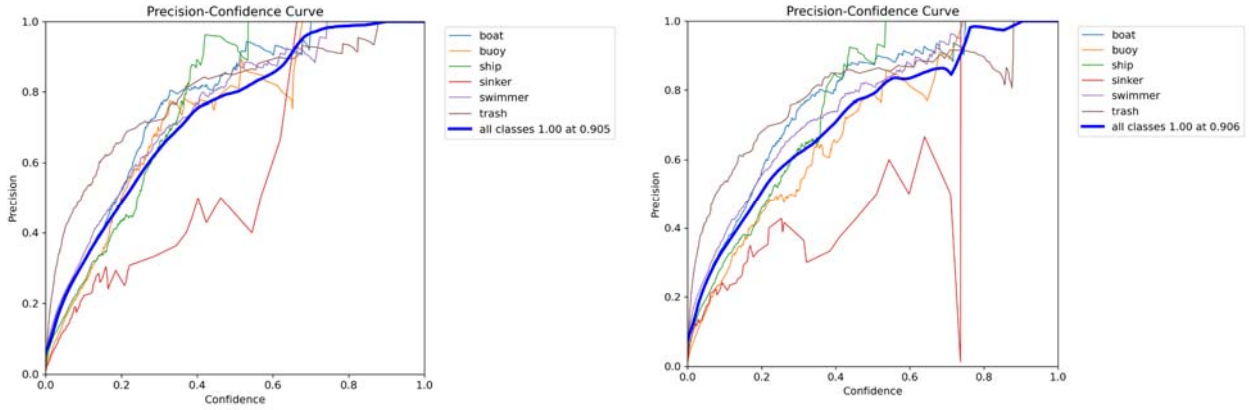
The generalized indices of precision and confidence for all varieties of the studied object classes have roughly equivalent values for YOLOv8 compared to YOLOv5 and are 0.906 versus 0.905, respectively.

The confusion matrices shown in Fig. 7 a and b, illustrate the percentage with which an object of a certain class will be determined (classified) by an object of another class for the YOLOv5 and YOLOv8 models, respectively.



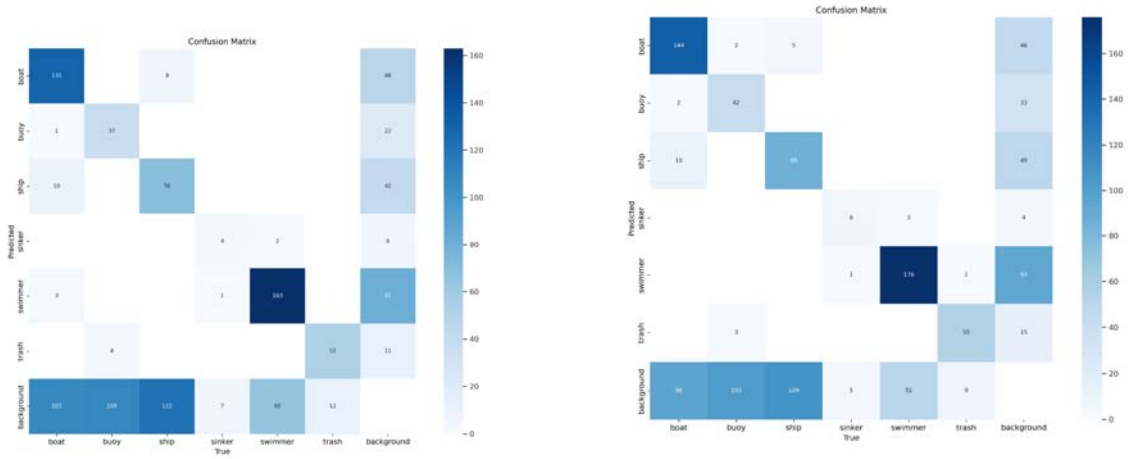

a b

Figure 5 – RC-metrics of models NN
a – YOLOv5; b – YOLOv8




a b

Figure 6 – PC-metrics of models NN
a – YOLOv5; b – YOLOv8




a b

Figure 7– Confusion matrix for NN
a – YOLOv5; b – YOLOv8

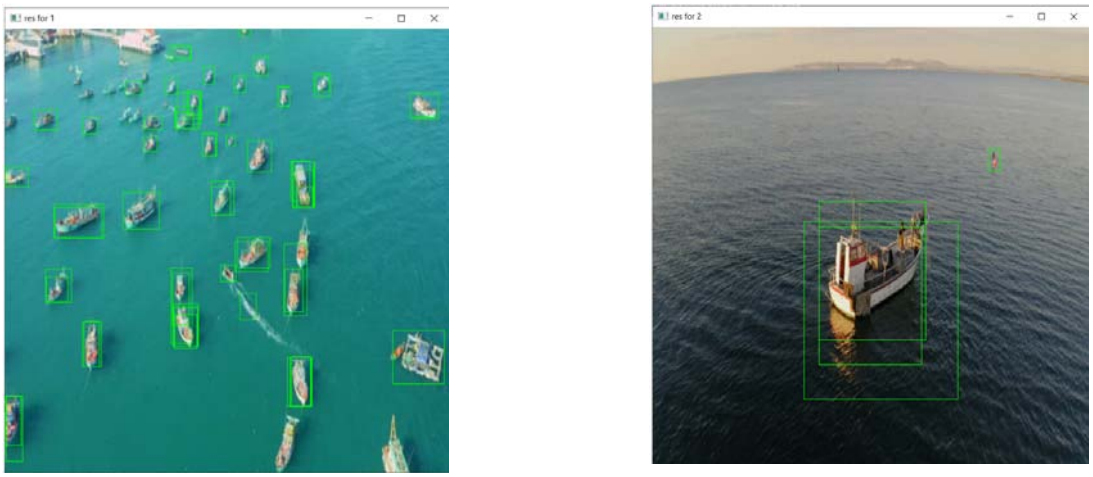

DOI 10.15588/1607-3274-2024-4-13

Examining in more detail the confusion matrices for the corresponding neural network models for each of the types of images recognized by objects, we get: the number of correctly recognized images of boats using YOLOv5 is significantly less than YOLOv8 and is 131 compared to 144 (9% difference); the number of correctly identified buoys is significantly different, 37 versus 42 (almost 12%); ships are recognized in the number of 70 to 86 (a large difference – about 19%); the number of images of drowning people is 4 compared to 6 (a very large difference of 33%), which is due to the small (unrepresentative) volume of the studied images; the number of photos of correctly recognized swimmers is 163 against 176 (7%); recognition of garbage images was performed with a slight difference of 53 vs. 55 (almost 4%), which leads to greater reliability and reliability of using the YOLOv8 neural network model compared to YOLOv5.

## 4 EXPERIMENTS

Examples below show classification process with reduced accuracy threshold down to 0.1 and with classification task dismissed. Recognized objects located correctly without explicit classification examples.

Test results depicted at Fig. 8.



a b

Figure 8 – Reduced accuracy threshold test results with classification function turned off
a – YOLOv5; b – YOLOv8

For test purposes, images were mostly marked correctly without exceptions dependent on object scale. It can be seen that model detects multiple objects at areas of actual objects. That type of behavior may be changed with implementation of bounding boxes filtering algorithm. Such algorithm reacts on a large cluster of overlapping detected objects, expanding single bound up to the borders of the biggest overlapping container so, grouping object cluster into one large object.

Selective object allocation pattern can be seen with objects being closer to the imaginary middle horizontal line being recognized better than others sticking closer to the top and bottom of the image. Such phenomenon can be explained with train dataset nature, where most of objects were placed at horizon level. In advance it can be seen that false object recognition is completely absent.

For the neural network, the training process of which is described in the corresponding section, it was necessary to write means for forming the on-board flight log and reporting on the found objects. As already mentioned in the analysis section, the Ultralytics API was chosen to interact with the model. It contains varitety of modules for dataset gathering, processing and filtering and some modules implement functionality connected with NN training. API depends on external libraries for some of its functions that installs automatically while using pip. Thus, the script that processes data in OpenCV image format looks like this:

```
import numpy as np
import cv2
from ultralytics import YOLO
yolo5_file = "./models/best5.pt"
yolo8_file = "./models/best8.pt"
under_test = YOLO(yolo5_file)
test_set =
[cv2.resize(cv2.imread(f"test_images/{i}.pn
g",cv2.IMREAD_COLOR),(640,640)) for for
i,data in enumerate(test_set):
r = under_test(data,imgsz = 640,conf=0.1)
print(r)
for s in r:
print(f"\n\nIMAGE #{i}")
boxes = s.boxes.xyxy.tolist()
for box in boxes:
box = list(map(int,box))
print(box)
test_set[i] =
cv2.rectangle(test_set[i],pt1=(box[0],box[1
]),pt2=(box[2],box[3]),color=(0,0,255),thic
kness=cv2.imshow('hello',test_set[i])
cv2.waitKey(0)
# closing all open windows
cv2.destroyAllWindows()
```

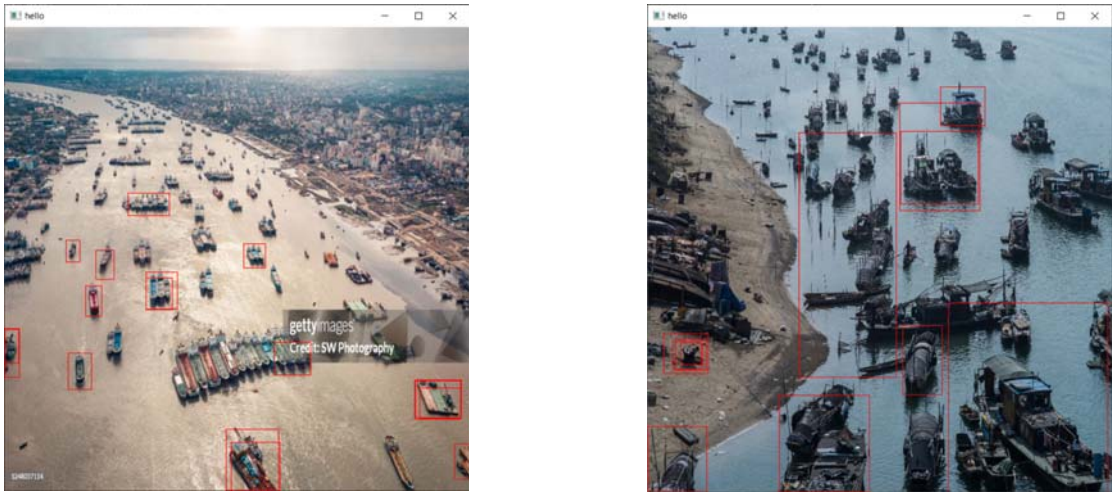

DOI 10.15588/1607-3274-2024-4-13

Fig. 9 a and b show images obtained after neural network processing. Certain objects are surrounded by several rectangles. This occurs due to the lowering of the accuracy threshold, as a result of which one object is classified into different classes. This is usually remedied by better training, but for the reasons mentioned earlier, another solution must be devised. This can be the implementation of a filter for post-processing of predictions of a neural network that will expand the boundaries of an existing rectangle in the case of a sufficient percentage of overlapping planes or create a new one in another case.

Fig. 10 a and b illustrate the uncertainty of the model regarding object ownership. We can see how due to the uncertainty of the model regarding the belonging of the object to a certain class, double allocation of objects appears. This can be avoided in two ways: retrain the model on a new dataset (which is ineffective for the given conditions of solving the problem) or write a filter that will process the selected objects and combine rectangles in an elementary way.

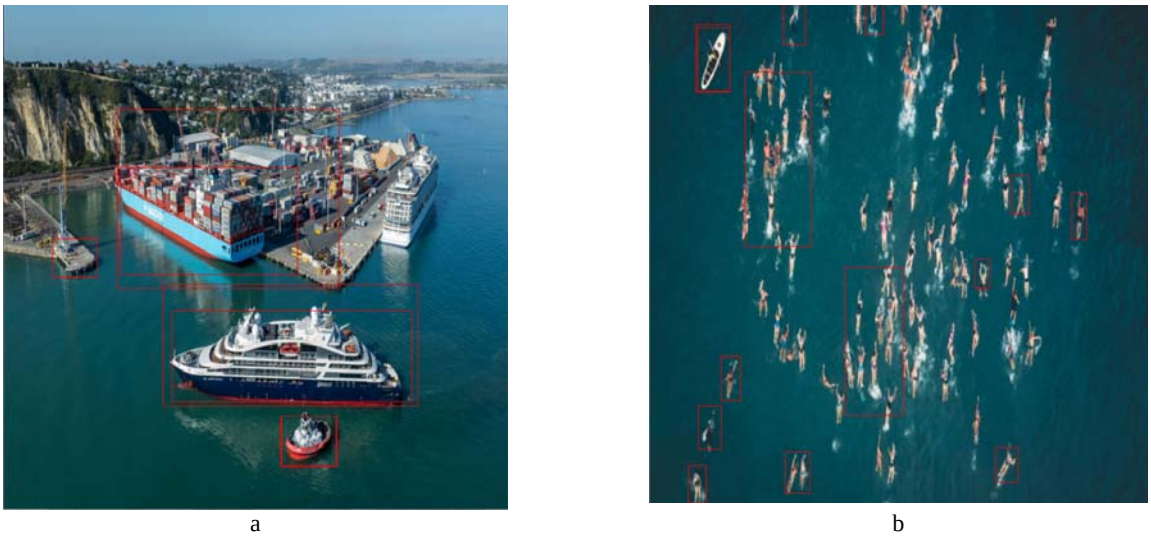
The specified block diagram of the algorithm performs the functions of filtering objects obtained after recognition by the neural network, namely, it analyzes the presence of an intersection of the limiting border in the form of a rectangle with a similar one for the case of the inability of the neural network to qualitatively classify this object.

Fig. 11 The block diagram of the algorithm shown in the figure 25 performs the functions of filtering objects obtained after recognition by the neural network, namely, it analyzes the presence of an intersection of the limiting border in the form of a rectangle with a similar one for the case of the inability of the neural network to qualitatively classify this object.




a b

Figure 9 – Images obtained after neural network processing



a b

Figure 10 – An illustration of the model's uncertainty about whether an object belongs to a certain class


DOI 10.15588/1607-3274-2024-4-13

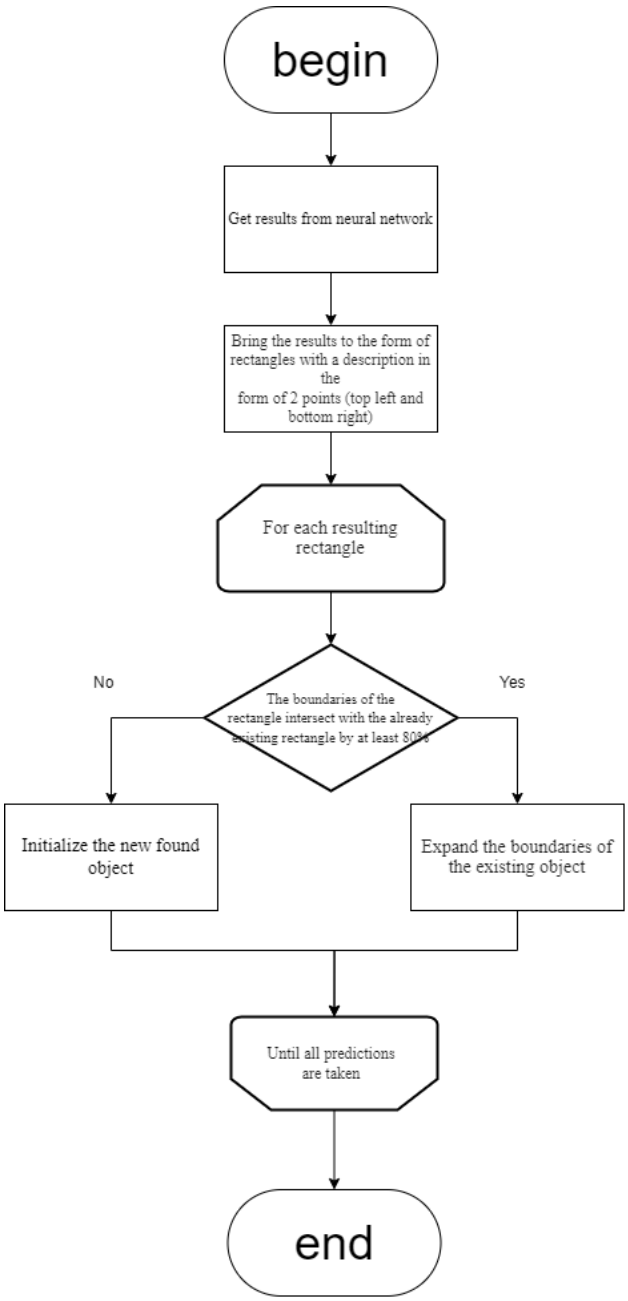
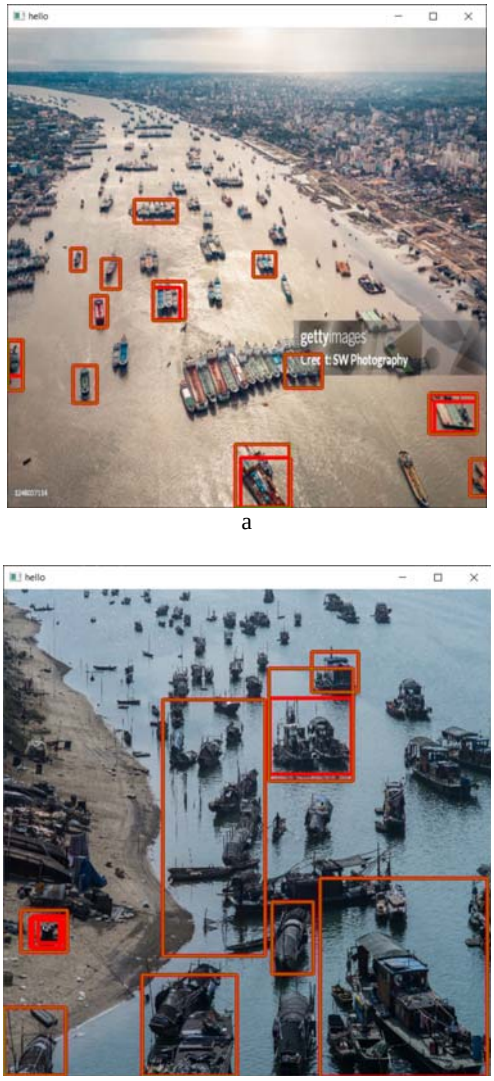



Figure 11 – Block diagram of the post-processing algorithm

The computational complexity of the algorithm was O(n). since the algorithm iterates through all received predictions, the memory complexity is O(log(n)) on average, O(n) in the worst case, because the resulting number of predictions either decreases or remains constant.

As a result of the implementation of the proposed algorithm, the clusters detected by the neural network are expanded. The set of rectangles in the previous step for certain objects is combined into one, thus we get rid of redundant predictions. The specified modification of the process of recognizing objects against the background of a reservoir, designed to eliminate redundancy in the prediction process, led to changes in the program script.




a

b

Figure 12 – Combining and filtering groupings from multiple predictions

As a result of the functioning of the proposed script, the number of recognized images was reduced. Fig. 12 a and b show how groups from several predictions are combined and filtered. Unprocessed predictions are marked in red, combined using the proposed algorithm and implemented using the created script are in green.

Original predictions are marked in red, post-processed ones in green. Test results reuire storage in some kind of logging system in order to perform system evaluation tests these. TinyDb – built-in Python module was used to secure image pathes in a single file. Mentioned tiny-db module does not support image storage unlike more complex systems such as PostgreSQL but this limitation may be avoided by storing file path in log file separately from the real extracted image. Limitations, provided in AI training paragraph, tell that original images also must be stored since such data may be used for gathering new dataset.

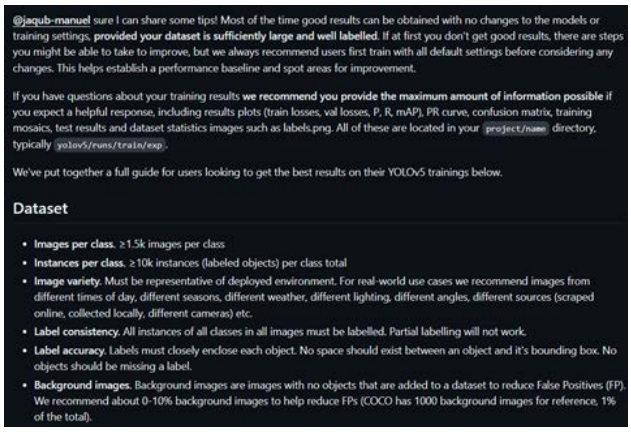
The general image processing algorithm was adjusted by implementing a filter to postprocess neural network


DOI 10.15588/1607-3274-2024-4-13

predictions, which improved prediction results by reducing the number of duplicate recognitions of the same objects. Scripts are offered that implement the specified post-processing and allow you to fix log directories of recognized objects and input images.

## 5 RESULTS

Gathering all previously said we can come to a conclusion objects detection part of the system works fine with exception of classification part being completely non-functional. Explanation for this involves dataset architecture, data sources, labelling process. These issues are common among API users and one of authors commented on ways of solving them. It is said that gathering bigger dataset with relevant data and proper labeling may reduce rate of misclassification cases.




Described system can be used with few extensions: as a originally designed system without function of object classification and also as a system for gathering relevant data for further implementation in a form of new train dataset since the current dataset contains of images taken from sailor boat and cannot perfectly fit goals of UAV reconnaissance.

A block diagram of the algorithm and corresponding scripts are proposed, which perform the functions of filtering the objects obtained after recognition by the neural network, namely, analyze the presence of an intersection of the limiting border in the form of a rectangle with a similar one for the case of the inability of the neural network to qualitatively classify this object. As a result of the implementation of the proposed algorithm, the clusters detected by the neural network are expanded. The set of rectangles in the previous step for certain objects is combined into one, thus we get rid of redundant predictions. For dataset gathering purposes logging subsystem with usage of tinydb was introduced.

The general image processing algorithm was adjusted by implementing a filter to postprocess neural network predictions, which improved prediction results by reducing the number of duplicate recognitions of the same objects. Scripts are offered that implement the specified post-processing and allow you to fix log directories of recognized objects and input images.


DOI 10.15588/1607-3274-2024-4-13

## CONCLUSIONS

The issue of reoccuring object recognition, caused by lowering the neural network's accuracy threshold, has been addressed. An algorithm for filtering predictions generated by a neural network, which excludes duplication of predictions, was proposed and tested, and a cut-off threshold was determined at which the object would be detected, but there would be no duplication. This achieves preservation of detection quality while eliminating duplication.

**The scientific novelty** consists in adding post-processing stage to NN predictions pipeline, which is based on an algorithm similar to the filtering algorithm already implemented in NN, which supresses NN-provided reoccuring object detections with provided algorithm based on image area intersection.

**The practical significance** of obtained results is implementation of post-processing algorithm and the created script based on it, a result was obtained where groups from several predictions are combined and filtered. As a result, these findings are saved in the database in the form they were obtained.

**Prospects for further research** are to improve the system by using UAVs as mobile cameras that will send video data to a central station where a sophisticated detector trained on relevant data will be deployed. This will make it possible to reduce the load on the UAV (both physical and informational), as well as reduce the cost of the UAV by the cost of the platform for deploying AI and move from the perspective of considering the UAV as a direct executor of the mission to the perspective of using it as a manipulator for the human-system AI.

## ACKNOWLEDGEMENTS

This article was made possible thanks to Oleksandr Rolik and Mykola Shynkevych, staff members of the Department of Information Systems and Technologies at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

## REFERENCES

1. Schunfelder P., Stebel F., Andreou N., Kunig M. Deep learning-based text detection and recognition on architectural floor plans, *Automation in Construction*, 2024, № 157. pp. 105–156. DOI: 10. 1016/j.autcon.2023.105156 (date of access: 06.05.2020)..
2. Giakoumoglou N., Pechlivani E. M., Tzovaras D. Generate-paste-blend-detect: Synthetic dataset for object detection in the agriculture domain, S*mart Agricultural Technology*, 2023, №5, pp. 100–105. DOI:10.1016/j.atech.2023.100258.
3. Ashourpour M., Azizpour G., Johansen K. Real-time defect and object detection in assembly line: A case for in-line quality inspection, *Lecture Notes in Mechanical Engineering*, 2024, pp. 99–106. DOI: 10.1007/978-3-031-38241-3_12.
4. Azevedo P., Santos V. Comparative analysis of multiple yolo-based target detectors and trackers for adas in

edge devices, *Robotics and Autonomous Systems*, 2024, № 171, pp. 87–95. DOI: 10.1016/j.robot.2023.104558.
5. Sanjai Siddharthan M., Aravind S., Sountharrajan S. Real-time road hazard classification using object detection with deep learning, *Lecture Notes in Networks and Systems*, 2024, № 789, LNNS, pp. 479– 492. DOI: 10.1007/978-981-99-6586-1_33.
6. Wei Z., Zhang Y., Wang X., Zhou J., Dou F., Xia Y. A yolov8-based approach for steel plate surface defect detection, *Metalurgija*, 2024, № 63, pp. 28–30. DOI: 10.3390/app14125325
7. Wu F., Zhang Y., Wang L., Hu Q., Fan S., Cai W. A deep learning-based lightweight model for the detection of marine fishes, *Journal of Marine Science and Engineering*, 2023, №11, pp. 19–34. DOI: 10.3390/jmse11112156.
8. Zhang G., Tang Y., Tang H., Li W., Wang L. A global lightweight deep learning model for express package detection, *Journal of Intelligent and Fuzzy Systems*, 2023, № 45. pp. 12013–12025. DOI: 10.3233/JIFS-232874.
9. Wang J., Dai H., Chen T., Liu H., Zhang X., Zhong Q., Lu R. Toward surface defect detection in electronics manufacturing by an accurate and lightweight yolo-style object detector, *Scientific Reports*, 2023, № 13, pp. 7–26. DOI: 10.1038/s41598-023-33804-w.
10. Li A., Zhang Z., Sun S., Feng M., Wu C. Multinet-gs: Structured road perception model based on multi-task convolutional neural network, *Electronics (Switzerland)*, 2023, № 12 (19), pp. 39–59. DOI: 10.3390/electronics12193994.
11. Han L., Ma C., Liu Y., Jia J., Sun J. Sc-yolov8: A security check model for the inspection of prohibited items in x-ray images, *Electronics (Switzerland)*, 2023, № 12 (20), pp. 208–223. DOI: 10.3390/electronics12204208.
12. Mao J., Wang L., Wang N., Hu Y., Sheng W. A novel method of human identification based on dental impression image, *Pattern Recognition*, 2023, № 144, pp. 109–126. DOI: 10.1016/j.patcog.2023.109864.
13. Kara E., Zhang G., Williams J. J., Ferrandez-Quinto G., Rhoden L. J., Kim M., Kutz J. N., Rahman A. Deep learning based object tracking in walking droplet and granular intruder experiments, *Journal of Real-Time Image Processing*, 2023, № 20, pp. 269–311 DOI: 10.1007/s11554-023-01341-4.
14. Zhou S., Zhong M., Chai X., Zhang N., Zhang Y., Sun Q., Sun T. Framework of rod-like crops sorting based on multi-object oriented detection and analysis, *Computers and Electronics in Agriculture*, 2024, № 216, pp. 108–145. DOI: 10.1016/j.compag.2023.108516
15. Shan P., Yang R., Xiao H., Zhang L., Liu Y., Fu Q., Zhao Y. Uavpnet: A balanced and enhanced uav object detection and pose recognition network, *Measurement: Journal of the International Measurement Confederation*, 2023, № 222, pp. 113–132. DOI: 10.1016/j.measurement.2023.113654.
16. Talaat F. M., ZainEldin H. An improved fire detection approach based on yolov8 for smart cities, *Neural Computing and Applications*, 2023, № 35, pp. 20939–20954. DOI: 10.1007/s00521-023-08809-1.
17. Liu S., Fan Q., Zhao C., Li S. Rtad: A real-time animal object detection model based on a large selective kernel and channel pruning, *Information (Switzerland)*, 2023, № 14 (10), pp. 535–550 DOI: 10.3390/info14100535.
18. Smolij V. About features of management preproduction of electronic vehicles, *Problems of Modeling and Design Automatization*, 2019, № 11, pp. 33–42. DOI: 10.31474/2074-7888-2019-1-33-42.
19. Su Y., Tan W., Dong Y., Xu W., Huang P., Zhang J., Zhang D. Enhancing concealed object detection in active millimeter wave images using wavelet transform, *Signal Processing*, 2024, № 216, pp. 109–122. DOI: 10.1016/j.sigpro.2023.109303.
20. Liu C., Wang K., Li Q., Zhao F., Zhao K., Ma H. Powerful-iou: More straightforward and faster bounding box regression loss with a nonmonotonic focusing mechanism, *Neural Networks*, 2024, № 170, pp. 276–284. DOI: 10.1016/j.neunet.2023.11.041.
21. Xu W., Liu C., Wang G., Zhao Y., Yu J., Muhammad A., Li D. Behavioral response of fish under ammonia nitrogen stress based on machine vision, *Engineering Applications of Artificial Intelligence*, 2024, № 128, pp. 107–134. DOI: 10.1016/j.engappai.2023.107442.
22. Dimauro G., Barbaro N., Camporeale M. G., Fiore V., Gelardi M., Scalera M. Deepcilia: Automated, deep-learning based engine for precise ciliary beat frequency estimation, *Biomedical Signal Processing and Control*, 2024, № 90, pp. 1002–1019. DOI: 10.1016/j.bspc.2023.105808.
23. Zhao X., Song Y. Improved ship detection with yolov8 enhanced with mobilevit and Gsconv, *Electronics (Switzerland)*, 2023, № 12(22), pp. 46–66. DOI: 10.3390/electronics12224666.
24. Smolij V. M., Smolij N. V., Sayapin S. P. Search and classification of objects in the zone of reservoirs and coastal zones, *CEUR Workshop Proceedings*, 2024, No. 3666, pp. 37–51. EID: 2-s2.0-85191443231
25. Hui J. mAP (mean Average Precision) for Object Detection. Medium. URL: https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173 (date of access: 15.06.2024).

Received 25.06.2024.
Accepted 28.10.2024.


DOI 10.15588/1607-3274-2024-4-13

УДК 681.325

# ПОСТ-ОБРОБКА ПРОГНОЗІВ ДЛЯ ПОЛІПШЕННЯ ЯКОСТІ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ПОВЕРХНІ ВОДОЙМ

**Смолій В. М.** – д-р техн. наук, професор, професор кафедри інформаційних систем та технологій Національного університету біоресурсів і природокористування України, Київ, Україна.

**Смолій Н. В.** – студент кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.

**Мокрієв М.В.** – канд. екон. наук, доцент, доцент кафедри інформаційних систем та технологій Національного університету біоресурсів і природокористування України, Київ, Україна.

## АНОТАЦІЯ

**Актуальність роботи** обумовлена масштабним поширенням технологій штучного інтелекту на процес виявлення і детекції об'єктів на поверхні водойм за допомогою БПЛА. Сучасні потреби в моніторингу водойм, особливо в контексті екологічного нагляду, охорони та управління ресурсами, вимагають точних і надійних рішень. Ця робота демонструє методи покращення роботи нейронних мереж і пропонує підходи до обробки передбачень НМ, навіть якщо вони натреновані на нерелевантних даних, що підвищує універсальність і ефективність технології.

**Мета роботи** – вирішення проблеми помилкового розпізнавання об'єктів на поверхні водойм, що обумовлено зменшенням порогу точності для нейронної мережі. Це забезпечує точнішу і надійнішу детекцію, зменшуючи кількість хибнопозитивних передбачень і підвищуючи ефективність системи загалом.

**Метод.** Запропоновано додавання етапу постобробки передбачень НМ, який базується на алгоритмі, схожому на вже впроваджений в НМ алгоритм фільтрації. Цей алгоритм пригнічує повторне виявлення об'єкта мережею та спирається на площі перетину виявлених прямокутників. Він використовує порогове значення 0.8 для двох точок прямокутника, що дозволяє ефективно знижувати кількість повторних передбачень та покращувати точність.

**Результати.** В результаті впровадження запропонованого алгоритму і створеного на його основі скрипта було досягнуто результат, при якому угруповання з кількох передбачень поєднуються та фільтруються. Отримані дані зберігаються в базі даних як знайдені та детектовані об'єкти. Запропонований алгоритм постобробки ефективно усуває надлишкові передбачення, при цьому зберігаючи точність прогнозу. Це забезпечує надійність системи і підвищує її продуктивність в реальних умовах.

**Висновки.** Детектовані зображення об'єктів на поверхні водойм зберігаються в базі даних у вигляді записів з унікальними ідентифікаторами файлових імен. Завдяки запропонованому алгоритму постобробки зображень формуються журнали для відпрацьованої місії з використанням тестових зображень, що забезпечує виключення можливості дублювання інформації. Це підвищує ефективність та надійність системи моніторингу, забезпечуючи точну і своєчасну детекцію об'єктів на поверхні водойм.

**КЛЮЧОВІ СЛОВА:** БПЛА, детекція, розпізнавані об'єкти, поверхня водойм, нейронна мережа, набір даних, модель, розподіл зображень, матриця помилок, навчальні метрики, збільшення, мозаїчне розміщення зображень, постобробка зображень, скрипт реалізації, журнал відпрацьованої місії, база даних.

## ЛІТЕРАТУРА

1. Deep learning-based text detection and recognition on architectural floor plans / [P. Schunfelder, F. Stebel, N. Andreou, M. Kunig] // Automation in Construction. – 2024. – № 157. – P. 105–156. DOI: 10. 1016/j.autcon.2023.105156 (date of access: 06.05.2020)..
2. Giakoumoglou N. Generate-paste-blend-detect: Synthetic dataset for object detection in the agriculture domain / N. Giakoumoglou, E. M. Pechlivani, D. Tzovaras // Smart Agricultural Technology. – 2023. – № 5. – P. 100–105. DOI: 10.1016/j.atech.2023.100258.
3. Ashourpour M. Real-time defect and object detection in assembly line: A case for in-line quality inspection / M. Ashourpour, G. Azizpour, K. Johansen // Lecture Notes in Mechanical Engineering. – 2024. – P. 99–106. DOI: 10.1007/978-3-031-38241-3_12.
4. Azevedo P. Comparative analysis of multiple yolo-based target detectors and trackers for adas in edge devices / P. Azevedo, V. Santos // Robotics and Autonomous Systems. – 2024. – № 171. – P. 87–95. DOI: 10. 1016/j.robot.2023.104558.
5. Sanjai Siddharthan M. Real-time road hazard classification using object detection with deep learning/ M. Sanjai Siddharthan, S. Aravind, S. Sountharrajan // Lecture Notes in Networks and Systems. – 2024. – № 789. LNNS. – P. 479–492. DOI: 10.1007/978-981-99-6586-1_33.
6. A yolov8-based approach for steel plate surface defect detection / [Z. Wei, Y. Zhang, X. Wang et al.] // Metalurgija. – 2024. – № 63. – P. 28–30. DOI: 10.3390/app14125325
7. A deep learning-based lightweight model for the detection of marine fishes/ [F. Wu, Y. Zhang, L. Wang et al.] // Journal of Marine Science and Engineering. – 2023. – № 11. – P. 19–34. DOI: 10.3390/jmse11112156.
8. A global lightweight deep learning model for express package detection / [G. Zhang, Y. Tang, H. Tang et al.] // Journal of Intelligent and Fuzzy Systems. – 2023. – № 45. – P. 12013–12025. DOI: 10.3233/JIFS-232874.
9. Toward surface defect detection in electronics manufacturing by an accurate and lightweight yolo-style object detector/ [J. Wang, H. Dai, T. Chen et al.] // Sci-


DOI 10.15588/1607-3274-2024-4-13

entific Reports. – 2023. – № 13. – P. 7–26. DOI: 10.1038/s41598-023-33804-w.
10. Multinet-gs:Structured road perception model based on multi-task convolutional neural network/ [A. Li, Z. Zhang, S. Sun et al.] // Electronics (Switzerland). – 2023. – № 12 (19). – P. 39–59. DOI: 10.3390/electronics12193994.
11. Sc-yolov8: A security check model for the inspection of prohibited items in x-ray images/ [L. Han, C. Ma, Y. Liu et al.] // Electronics (Switzerland). – 2023. – № 12 (20). – P. 208–223. DOI: 10.3390/electronics12204208.
12. A novel method of human identification based on dental impression image/ [J. Mao, L. Wang, N. Wang et al.] //Pattern Recognition. – 2023. – № 144. – P. 109 – 126. DOI: 10.1016/j.patcog.2023.109864.
13. Deep learning based object tracking in walking droplet and granular intruder experiments/ [E. Kara, G. Zhang, J. J. Williams et al.] // Journal of Real-Time Image Processing. – 2023.– № 20.– P. 269–311 DOI: 10.1007/s11554-023-01341-4.
14. Framework of rod-like crops sorting based on multi-object oriented detection and analysis/ [S. Zhou, M. Zhong, X. Chai et al.] //Computers and Electronics in Agriculture. – 2024. – № 216. – P. 108–145. DOI: 10.1016/j.compag.2023.108516
15. Uavpnet: A balanced and enhanced uav object detection and pose recognition network/ [P. Shan, R. Yang, H. Xiao et al.] // Measurement: Journal of the International Measurement Confederation. – 2023. – № 222. – P. 113–132. DOI: 10.1016/j.measurement.2023.113654.
16. Talaat F. M. An improved fire detection approach based on yolov8 for smart cities/ F. M. Talaat, H. ZainEldin // Neural Computing and Applications. – 2023. – № 35. – P. 20939–20954. DOI: 10.1007/s00521-023-08809-1.
17. Rtad: A real-time animal object detection model based on a large selective kernel and channel pruning / [S. Liu, Q. Fan, C. Zhao, S. Li] // Information (Switzerland). – 2023. – № 14 (10). – P. 535–550. DOI: 10.3390/info14100535.
18. Smolij V. About features of management preproduction of electronic vehicles / V. Smolij // Problems of Modeling and Design Automatization. – 2019. – № 11. – P. 33–42. DOI: 10.31474/2074-7888-2019-1-33-42.
19. Enhancing concealed object detection in active millimeter wave images using wavelet transform / [Y. Su, W. Tan, Y. Dong et al.] // Signal Processing. – 2024. – № 216. – P. 109–122. DOI: 10.1016/j.sigpro.2023.109303.
20. Powerful-iou: More straightforward and faster bounding box regression loss with a nonmonotonic focusing mechanism/ [C. Liu, K. Wang, Q. Li et al.] // Neural Networks. – 2024. – № 170. – P. 276–284. DOI: 10.1016/j.neunet.2023.11.041.
21. Behavioral response of fish under ammonia nitrogen stress based on machine vision / [W. Xu, C. Liu, G. Wang et al.] // Engineering Applications of Artificial Intelligence. – 2024. – № 128. – P. 107–134. DOI: 10.1016/j.engappai.2023.107442.
22. Deepcilia: Automated, deep-learning based engine for precise ciliary beat frequency estimation/ [G. Dimauro, N. Barbaro, M. G. Camporeale et al.] // Biomedical Signal Processing and Control. – 2024. – № 90. – P. 1002–1019. DOI: 10.1016/j.bspc.2023.105808.
23. Zhao X. Improved ship detection with yolov8 enhanced with mobilevit and Gsconv / X. Zhao, Y. Song // Electronics (Switzerland). – 2023. – № 12(22). – P. 46–66. DOI: 10.3390/electronics12224666.
24. Smolij V. M. Search and classification of objects in the zone of reservoirs and coastal zones / V. M. Smolij, N. V. Smolij, S. P. Sayapin // CEUR Workshop Proceedings – 2024. – No. 3666. – P. 37–51. EID: 2-s2.0-85191443231
25. Hui J. mAP (mean Average Precision) for Object Detection. Medium. URL: https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173 (date of access: 15.06.2024).


DOI 10.15588/1607-3274-2024-4-13

# ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

# PROGRESSIVE INFORMATION TECHNOLOGIES

UDC 004.94 : 004.2

## ALGORITHMIC DIFFERENCES OF COMPLETE AND PARTIAL ALGEBRAIC SYNTHESIS OF A FINITE STATE MACHINE WITH DATAPATH OF TRANSITIONS

**Babakov R. M.** – Dr. Sc., Associate Professor, Professor of the Information Technologies Department, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine.

**Barkalov A. A.** – Dr. Sc., Professor, Professor of the Institute of Computer Science and Electronics, University of Zielona Gora, Zielona Gora, Poland.

**Titarenko L. A.** – Dr. Sc., Professor, Professor of the Institute of Computer Science and Electronics, University of Zielona Gora, Zielona Gora, Poland.

**Voitenko M. O.** – Student of the Faculty of Information and Applied Technologies, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine.

**ABSTRACT**

**Context.** The problem of algorithmization the search for formal solutions of the problem of algebraic synthesis of a finite state ma-chine with datapath of transitions is considered. The concept of complete and partial solutions of this problem is proposed. The object of research is the automated synthesis of the finite state machine in the part of the function of transitions without taking into account the function of outputs. The basis of the algebraic implementation of the transition function is the author's approach to the transformation of state codes using a set of arithmetic and logical operations. The search for formal solutions to the problem of algebraic synthesis is a complex process that requires the use of appropriate methods and algorithms aimed at special coding of states and mapping of operations of transitions to individual state machine transitions. The use of partial solutions of the problem of algebraic synthesis can contribute to reducing the executing time of such algorithms and reducing the overall design time of digital control devices based on a finite state machine with an datapath of transitions.

**Objective.** Development and research of algorithms for finding complete and partial solutions to the problem of algebraic synthesis of a finite state machine with datapath of transitions.

**Method.** The research is based on the structure of a finite state machine with datapath of transitions. The synthesis of the circuit of the state machine involves the preliminary solution of the problem of algebraic synthesis. The result is the so-called formal solution of this problem, which contains two components. The first component is defined state codes, the second component is arithmetic and logic operations mapped to separate state machine transitions. Finite state machine can be synthesized in that case if transformation of given state codes in the process of making transitions is possible using a given set of operations. Verification of this possibility is carried out using the known matrix approach. It involves the formation and element-by-element mapping of two matrices – the matrix of transitions and the combined matrix of operations. As a result, a coverage matrix is formed, which reflects the possibility of implementing (covering) of state machine transitions by specified arithmetic and logic operations. Changing the way of encoding states or the set of operations can give different solutions to the problem of algebraic synthesis with a greater or lesser number of covered state machine transitions.

**Results.** Using the example of an abstract graph-scheme of the control algorithm, it is demonstrated that the solution of the problem of algebraic synthesis of a finite state machine with datapath of transitions can be considered a situation when one or more state machine transitions cannot be implemented using a given set of operations. It is proposed to call such situation as a partial solution of the problem of algebraic synthesis. The implementation of all transitions by specified operations gives a complete solution of this problem, but the number of complete solutions is always much smaller than the number of partial solutions. Therefore, in the general case, the search for complete solutions takes much more time and, moreover, is not always possible.

**Conclusions.** The design of a logical circuit of a finite state machine with datapath of transitions is possible in the case of a complete or partial solution of the problem of algebraic synthesis. In the case of a partial solution, those transitions that cannot be implemented by any of the available operations are implemented in a canonical way using a system of Boolean equations. The search for complete solutions generally takes more time than the search for partial solutions. This makes actual the development of algorithms and methods of synthesis of this class of finite state machine, based on the search for partial solutions of the problem of algebraic synthesis.

**KEYWORDS:** finite state machine, datapath of transitions, graph-scheme of algorithm, arithmetic and logical operations, algebraic synthesis, partial solution.


DOI 10.15588/1607-3274-2024-4-14

## ABBREVIATIONS

FSM is a finite state machine;
DT is a datapath of transitions;
GSA is a graph-scheme of algorithm;
TO is a transitions operation.

## NOMENCLATURE

$A$ is a sets of FSM states;
$X$ is a sets of logical conditions;
$Y$ is a sets of microoperations;
$M$ is a number of FSM states;
$L$ is a number of logical conditions analyzed by FSM;
$P$ is a number of microoperations formed by FSM;
$R$ is a bit capacity of state code;
$O$ is a set of transitions operations;
$O_i$ is an element of set $O$;
$I$ is a number of TO;
$B$ is a number of uncovered FSM transitions;
$B_{min}$ is a minimal number of uncovered transitions;
$T$ is a code of FSM current state;
$D$ is a code of next state;
$G$ is a graph-scheme of algorithm.

## INTRODUCTION

Digital systems take a significant place in various spheres of human activity [1]. An integral component of a digital system is a control unit, the function of which is to coordinate the operation of all elements of the system [2, 3]. One of the ways of implementing a control unit is a finite state machine (FSM), in which a given control algorithm is implemented in the form of a logic circuit [4-6]. Compared to other types of control units, FSM is characterized by maximum speed due to the possibility of realizing multidirectional automatic transitions in one operation cycle. The disadvantage of FSM circuit is hardware expenses, which are maximum compared to other types of control units [2, 3, 7]. This worsens such characteristics of the FSM circuit as cost, energy consumption, dimensions, reliability, etc. [8, 9]. In this aspect, the scientific and practical problem of reducing hardware expenses in the FSM logic circuit is actual [3, 9, 10].

One of the well-known approaches to reducing hardware expenses in the FSM circuit is so-called operational transformation of state codes [11]. According to it, the transformation of state codes during state machine transitions is carried out not according to the system of canonical Boolean equations, but with using of a given set of arithmetic and logical operations (transitions operations, TO), which are performed on the code of the current state of the finite state machine. A set of such operations is implemented in the form of a set of corresponding combinational circuits, which form the so-called datapath of transitions (DT). The use of DT in the structure of FSM gives rise to the structure of a finite state machine with datapath of transitions (FSM with DT) [11]. The reduction of hardware expenses in the FSM with DT circuit compared to the canonical FSM is achieved due to the fact that as part of the DT, each combinational circuit is able to implement any number of FSM transitions with fixed hardware expenses for their implementation.

The synthesis of FSM with DT involves two main stages: algebraic synthesis and synthesis of the logic circuit of the FSM based on the results of algebraic synthesis [12]. Algebraic synthesis of the FSM is currently not sufficiently formalized and is not presented in the form of a method or an algorithmic and software implementation. This complicates the process of automated design and narrows the area of application of this FSM class.

In [13], it is shown that the algebraic synthesis of FSM with DT is reduced to the solution of the problem of algebraic synthesis. The search for solutions of this problem is proposed to be carried out using the comparison of special matrices, the first of which reflects the system of transitions of the finite state machine with a given coding of states, and the second reflects the possibilities of converting state codes using a given set of arithmetic and logic operations. The rules for constructing and comparing such matrices are considered in [13].

**The object of the study** is the algebraic synthesis of a finite state machine with datapath of transitions.

Algebraic synthesis is considered completed under two conditions: if all states of the FSM are coded; if for each FSM transition, the transformation of the current state code into the transition state code can be performed using one of the specified transitions operations.

**The subject of the study** are algorithms for finding complete and partial solutions of the problem of algebraic synthesis of FSM with DT.

**The purpose of the work** is the development and comparative research of algorithms of finding complete and partial solutions of the problem of algebraic synthesis of finite state machine with datapath of transitions.

## 1 PROBLEM STATEMENT

Let the finite state machine with datapath of transitions be given in the form of a graph-scheme of the algorithm (GSA) [3, 10]. According to this GSA, a set of states $A=\{a_1, ..., a_M\}$, a set of input signals $X=\{x_1, ..., x_L\}$ and a set of microoperations $Y=\{y_1, ..., y_P\}$ are formed. The set of transitions operations $O = \{O_1, \ldots, O_I\}$ is also given. Each element $O_i \in O$ represents a certain arithmetic or logical operation, which provides an appropriate interpretation of the binary codes of the FSM states. The GSA and the set $O$ are input data for the algebraic synthesis of FSM with DT.

This article solves the problem of developing algorithms for obtaining complete and partial solutions of problems of algebraic synthesis of FSM with DT for given GSA and a set of transition operations.

## 2 REVIEW OF THE LITERATURE

Known methods of optimizing the logic circuit of a finite state machine usually lead to changes in its structure [3, 6, 10]. This article considers the structure of finite state machine with datapath of transitions [11]. In it, the


DOI 10.15588/1607-3274-2024-4-14

transition function is implemented on the basis of the datapath of transitions, which provides a set of arithmetic or logical operations for the transformation of state codes. The synthesis of FSM with DT is considered in [12].

Let the FSM with DT be given by GSA *G* (Fig. 1). The operation vertices of this GSA do not contain microoperations and are shown empty. This is done in order to focus precisely on the function of transitions of the FSM, and not on the function of outputs. The output function in FSM with DT is implemented in the same way as in FSM with a canonical structure [2–5, 10] and is not considered in this paper.

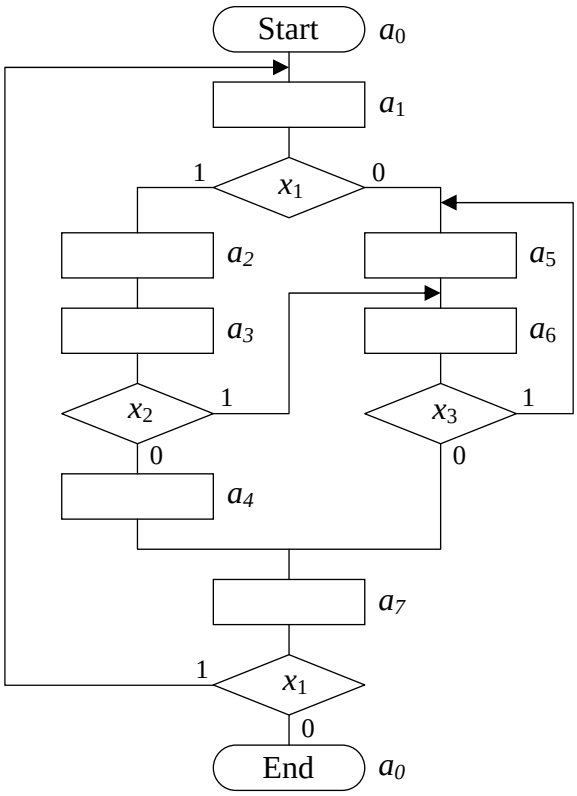



Figure 1 – Graph-scheme of algorithm *G*

GSA *G* contains 8 states of the Moore state machine [10] $a_0 - a_7$, for coding of which $R = 3$ binary digits are sufficient. GSA contains both conditional and unconditional FSM transitions, the total number of which is 12.

Let the set of TO $O = \{O_1, O_2, O_3\}$ be also given, where

$$O_1: D = T + 1_{10}; \quad (1)$$
$$O_2: D = T \vee 010_2; \quad (2)$$
$$O_3: D = T \oplus 110_2. \quad (3)$$

In expressions (1)–(3), the symbol *T* means code of current state, the symbol *D* means code of transition state. TO $O_1$ represents the addition of one, TO $O_2$ is a disjunction with binary constant 010, TO $O_3$ is XOR operation with binary constant 110.

The result of any TO is code of transition state of the FSM. This means that result of TO must always have the same bit size as the argument of TO (current state code). This bit size in the article is denoted by the symbol *R* and for GSA *G* is equal to 3. If TO $O_2$ and $O_3$ are bitwise logical operations, TO $O_1$ is an arithmetic operation (increment). If result of the increment is one digit greater than the argument, *R* lower digits should be taken as the result of the TO. Therefore, it is more correct to write TO $O_1$ as follows:

$$O_1: D = (T + 1_{10}) \bmod 8. \quad (4)$$

Here, *mod* means *modulo* – the operation of taking the remainder of a devision the value $(T + 1_{10})$ by 8, where the number 8 is calculated as $2^R$. This is equivalent to taking three lower digits from the increment result.

All arithmetic and logical operations should be arranged in a similar way. For example, if the logical operation of a cyclic bit shift is used, the shift must be carried out within *R* bits, that is, within the bit size of the state code of the FSM. If the arithmetic operation of adding the decimal constant 5 is used, its result is also taken modulo $2^R = 8$. That is, $(6 + 5) \bmod 8 = 3$.

Now consider the algebraic synthesis of FSM with DT. It consists in the coordinated execution of the following actions [12]:

1) each state of the FSM is assigned an unique code from a set of admissible state codes (usually from the range of integers $[0; 2^R–1]$);

2) each FSM transition is mapped to a certain operation from the set *O* (the same operation can be mapped to several transitions).

The result of these actions should be the operational implementation of all or the vast majority of state machine transitions. A transition has an operational implementation if the code of the initial state is transformed into the code of the state of the transition using the TO mapped to this transition. If none of the given TOs allows such a transformation, the transition has no operational implementation with the selected states coding and the given set of TOs.

In Fig. 2 an example of operational implementation of all transitions of GSA *G* for the set of TOs formed by operations (1)–(3) is shown.

In Fig. 2, the decimal code of state and its binary equivalent are recorded in each vertex marked by the state of the Moore FSM. Decimal codes should be considered when the transition to this state or from this state is carried out using an arithmetic TO. Binary codes are required in the case of using logical TO. Also in Fig. 2, each transition branch is marked by an operation mapped at this transition. For ease of understanding, TO $O_1$ is marked with "+1", TO $O_2$ – with "$\vee$ 010", TO $O_3$ – with "$\oplus$ 110".

It can be noted that in Fig. 2, each FSM transition is implemented (covered) by one of the given TOs. For example, a conditional transition from state $a_7$ with code $3_{10} = 011_2$ to state $a_1$ with code $5_{10} = 101_2$ is covered by the operation "$\oplus$ 110", since $011_2 \oplus 110_2 = 101_2$.


DOI 10.15588/1607-3274-2024-4-14

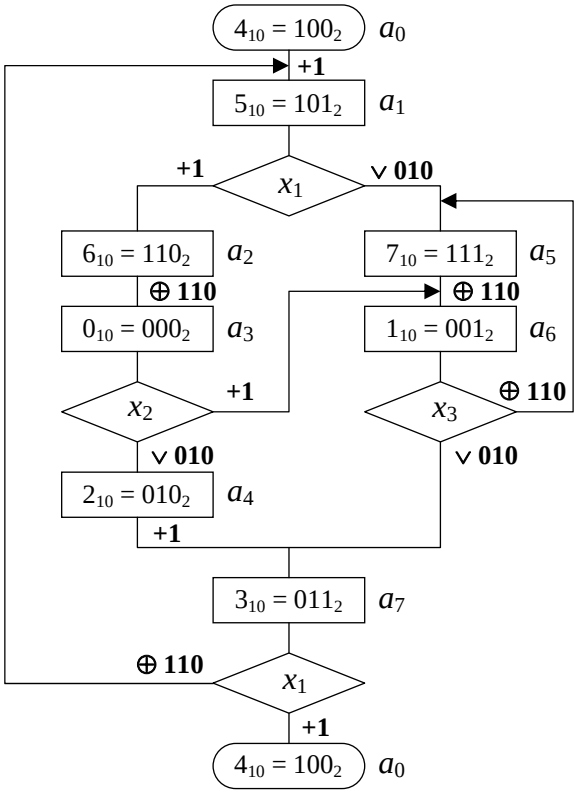



Figure 2 – Example of operational implementation of all transitions of GSA *G*

From Fig. 2, you can see that each TO is used to implement several FSM transitions. However, as part of the datapath of transitions, the circuit of each TO has fixed structure and fixed hardware expenses, which do not depend on the number of FSM transitions covered by this TO [11, 12]. For a large-sized GSA, the incrementer circuit can implement dozens and hundreds of transitions with fixed hardware expenses for their implementation. This is the basis of the well-known class of counter-based finite state machines [2]. FSM with DT is positioned as a generalization of an FSM with counter for the case of using a set of different arithmetic and logical operations. It is the possibility of implementing many FSM transitions with the help of a fixed set of TOs with fixed hardware expenses that ensures the saving of resources for the implementation the transition function of the FSM with DT compared to the canonical FSM [11].

If the states are coded in a different way or a different set of TOs is specified, part of the FSM transitions may remain uncovered. Consider a fragment of Fig. 2, in which codes of states of $a_2$ and $a_3$ are swapped (Fig. 3).

On this fragment, transitions $a_1 \to a_2$, $a_3 \to a_4$ and $a_3 \to a_6$ cannot be realized by any of the TOs $O_1 - O_3$. At the same time, the transition $a_2 \to a_3$ can be implemented using TO $O_3$. So, changing the codes of the two states added three FSM transitions, not covered by any of the given TOs.

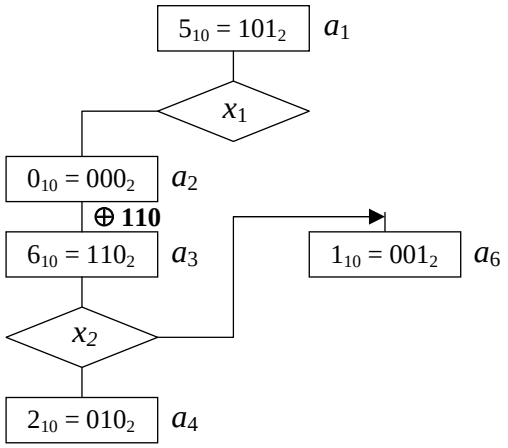



Figure 3 – Fragment of GSA *G* with uncovered state machine transitions

After carrying out the algebraic synthesis of FSM with DT, there is a stage of synthesis of the logical circuit of the FSM. This stage, in contrast to algebraic synthesis, is sufficiently formalized and is discussed in [12].

In [13], the situation when all FSM transitions are covered by given TOs is called a “formal solution to the problem of algebraic synthesis of FSM with DT”. The term “formal” here means that in the presence of such a solution, the functioning of the FSM is possible according to the principle of operational transformation of state codes and the synthesis of the logical circuit of the FSM is possible. The formal solution does not guarantee a reduction in the hardware expenses in the FSM circuit, but only ensures the functioning of the FSM based on the datapath of transitions.

The examples given in [13] demonstrate the possibility of the existence of a set of formal solutions of the problem of algebraic synthesis for a given GSA. Also in [13] the concepts of effective and optimal solutions are introduced. These solutions form proper subsets on the set of formal solutions. They are interesting in that their use makes it possible to obtain a gain in hardware expenses in the resulting circuit of the FSM with DT compared to other FSM structures.

It should be noted that the synthesis of the logical circuit of FSM with DT is possible even if part of FSM transitions remains uncovered by existing transitions operations [14]. In this case, all transitions not covered by TOs are implemented separately from other transitions according to the same principle as in the transition circuit of canonical FSM [2, 3, 10]. For them, a system of Boolean equations of the form (5) is formed, for which a corresponding combinational circuit is synthesized.

$$D = D(T, X). \qquad (5)$$

In expression (5), the transition function $D$ depends on the current state $T$ and input signals $X$. The number of system equations is equal to $R$ (bit size of the state code).

The combinational circuit synthesized in this way is considered as one of the operations of set $O$. It is added to the datapath of transitions and receives a separate code,


DOI 10.15588/1607-3274-2024-4-14

according to which its output is multiplexed with the outputs of other OPs circuits. For the considered example, the structure of the datapath is shown in Fig. 4. Blocks $O_1 – O_3$ correspond to transitions operations from the set $O$. Block $O_4$ contains a combinational circuit constructed according to equation (5). It can be considered that the $O_4$ block implements a non-standard logical TO, which is performed on the code of the current state $T$ and the input signals $X$. The synthesis of such block is discussed in detail in [14]. In Fig. 4, the multiplexer MUX under the control of the signals of the TO code $W$ forms the code of the next state $D$, which enters the memory register of the FSM (not shown in the figure).

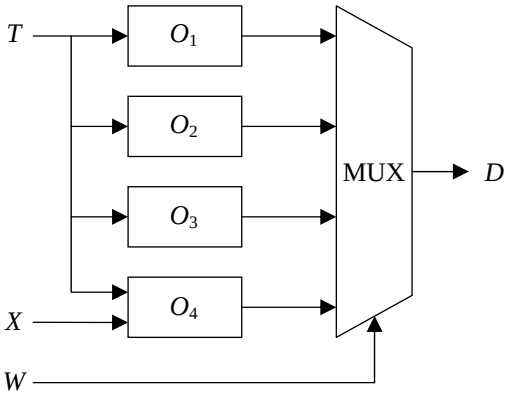



Figure 4 – Structure of datapath of transitions
in case of canonical realization of several transitions

## 3 MATERIALS AND METHODS

According to the results of algebraic synthesis, two situations are possible:

1. All FSM transitions are covered by the specified TOs.

2. Part of FSM transitions remained not covered by the given TOs and should be implemented in a canonical way according to the system of Boolean equations.

Let us clarify the meaning of a formal solution of the problem of algebraic synthesis of FSM with DT, given in [13]. We will consider the result of algebraic synthesis to be a formal solution, regardless of the number of covered transitions. To distinguish between full and partial coverage of transitions, we introduce the following definitions.

A solution in which all FSM transitions are covered by a given set of transitions operations will be called a *complete solution of the problem of algebraic synthesis of FSM with DT.*

*A partial solution of the problem of algebraic synthesis of FSM with DT* will be called a solution in which part of the transitions remains uncovered by a given set of TOs.

In a partial solution, the number of uncovered transitions can vary from 1 to the total number of all transitions of a given FSM. In the extreme case, if all transitions remain uncovered, the FSM with DT degenerates into a canonical FSM with the according structure [3, 10].

The purpose of differentiating complete and partial solutions is as follows.

Currently, there are no effective algorithms for finding complete solutions to the problem of algebraic synthesis of FSM with DT. It is not obvious what codes should be assigned to the FSM states so that all transitions be covered by a given set of operations. With a fixed set of TOs, the search for formal solutions can be reduced to enumeration of variants for coding states with a set of admissible codes. For each encoding variant, a transition coverage check must be performed.

During the enumeration of variants of states encoding, the complete solution can be found at the beginning of the enumeration, at the very end of the enumeration, or not found at all. At the same time, going through all possible options for coding states requires considerable time. If the GSA contains $M$ states, for the coding of which $R$ binary bits are enough, then the number of state coding variants $N$ is determined by expression (6).

$$N = \frac{(2^R)!}{(2^R - M)!}. \qquad (6)$$

For our example, with $R = 3$, $M = 8$, we have $N = 40320$ (taking into account $0! = 1$). However, for FSM with 10 states, we will get $N = 29 \cdot 10^9$ coding variants. For GSA with $M > 16$, the number of state coding variants does not allow considering all variants within reasonable time limits.

We will proceed from the assumption that all complete solutions will lead to the same or similar hardware expenses in the FSM circuit. Therefore, it is sufficient to find only one complete solution, after which the enumeration of state coding variants can be stopped. This allows us to propose an algorithm for finding a complete solution to the problem of algebraic synthesis of FSM with DT, the flowchart of which is shown in Fig. 5.

The peculiarity of this algorithm is that the search for a solution takes place until the first solution is found. At the same time, the algorithm does not guarantee finding a complete solution, since a complete solution may not exist for a given GSA and a given set of TOs. The following statements should be considered valid:

1. The probability of finding a complete solution increases with an increase in the number of transitions operations used.

2. The probability of finding a complete solution decreases with an increase in the number of FSM states with a fixed set of TO.

It follows from the first statement that the smaller the number of TOs used, the lower the probability of finding a complete solution. However, the economy of hardware expenses in the circuit of FSM with DT is achieved precisely due to the reduction of the number of TOs and the simplification of the circuit of datapath of transitions. Therefore, the desire to reduce the number of used TOs contributes to the failure to find complete solutions.

It follows from the second statement that increasing the complexity of the control algorithm implemented by FSM also does not contribute to finding complete solutions. The complexity of the algorithm in this case is


DOI 10.15588/1607-3274-2024-4-14

measured by the number of its states, which affects the bit size $R$ of state code. Therefore, the more complex the FSM is, the more difficult it is to cover all its transitions with a given set of TOs, since the number of state coding variants increases, according to (5), in a non-linear dependence.

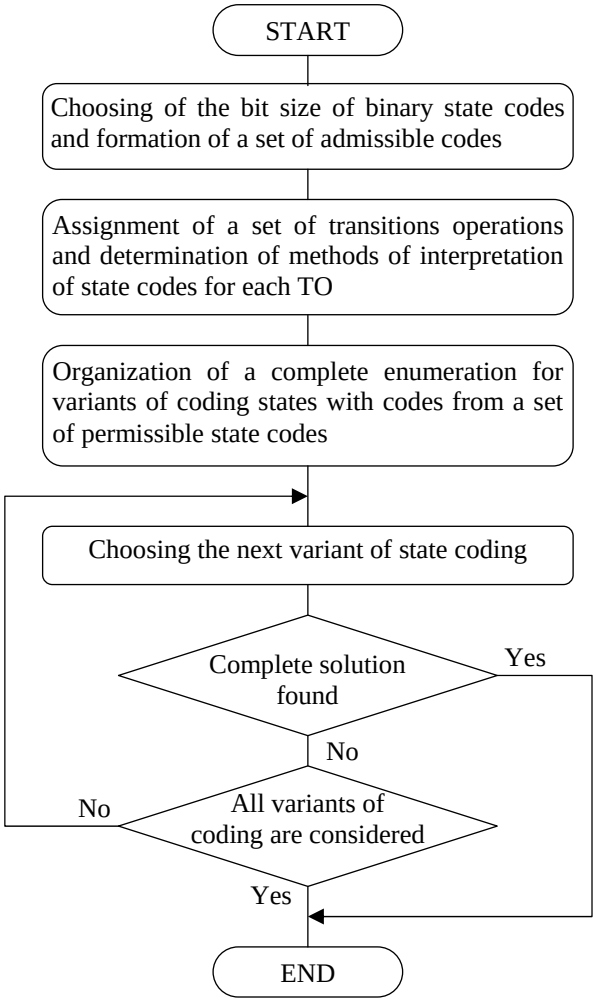



Figure 5 – Flowchart of algorithm of finding a complete solution of problem of algebraic synthesis for given set of transitions operations

Thus, the desire to reduce the number of TOs and the increasing complexity of control algorithms decrease the probability and increase the time of finding complete solutions to the problem of algebraic synthesis. The algorithm shown in Fig. 5 should be used only if, for some reason, it is inadmissible to have uncovered transitions and to implement them in a canonical way using a system of Boolean equations. The following can be considered as such reasons:

– the task was set to reduce hardware expenses in the FSM with DT to the minimum possible level;

– a pre-synthesized device is used as an DT, which cannot be changed.

If the algorithm for finding a complete solution does not find a result, the algebraic synthesis of FSM with DT under the given conditions becomes impossible. To eliminate this shortcoming, an algorithm for finding a partial solution is proposed, the generalized flowchart of which is shown in Fig. 6.

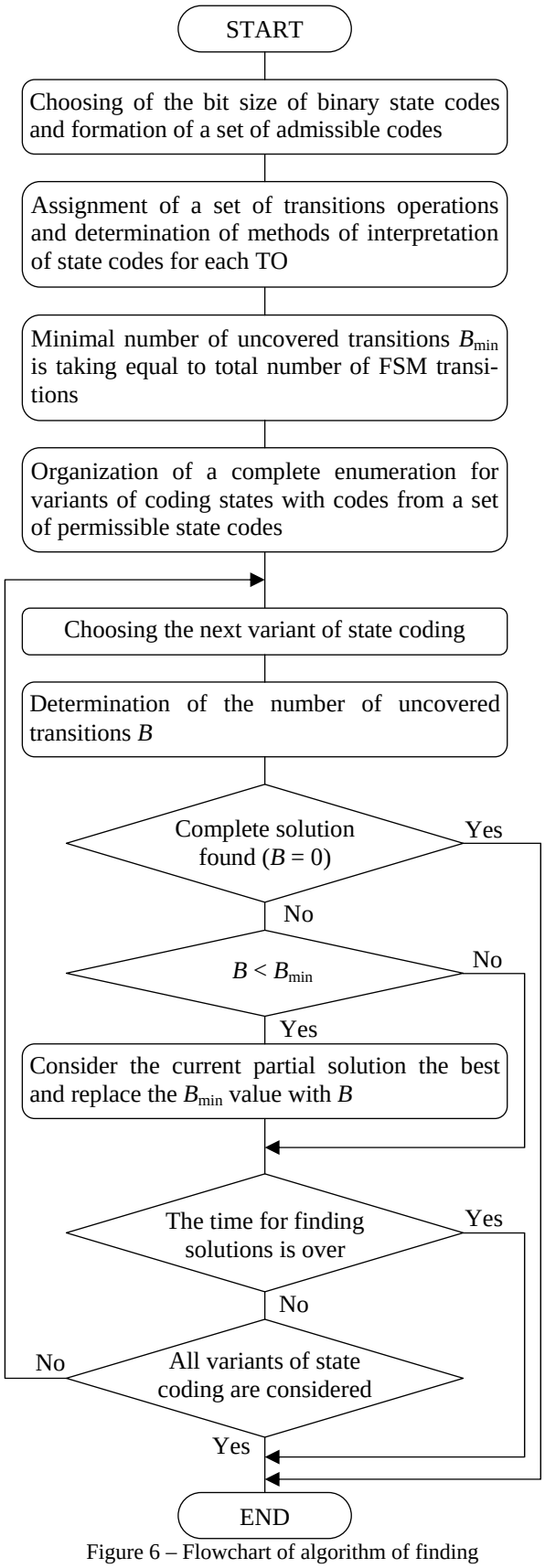



Figure 6 – Flowchart of algorithm of finding a partial solution of problem of algebraic synthesis


DOI 10.15588/1607-3274-2024-4-14

The differences between this algorithm and the algorithm for finding a complete solution are as follows.

1. This algorithm always gives a result sufficient for the synthesis of the logic circuit of the FSM. This result can be either a complete solution or a partial solution, but it will always be found. The algorithm for finding a complete solution (Fig. 5) may not give a result at all, which will make algebraic and further synthesis of the FSM impossible.

2. If the time constraints are set, the algorithm in the given time finds among the considered partial solutions one that has the minimum number of uncovered transitions. This makes it possible to maximally simplify the circuit that corresponds to the canonical implementation of uncovered transitions.

3. If time constraints are not specified, the algorithm after enumeration through all possible states coding variants gives the best possible result for a given GSA with a given set of transitions operations. It can be either a complete or a partial solution, which can be considered optimal under the given conditions of algebraic synthesis.

As in the previous algorithm, when developing the algorithm for finding partial solutions, the assumption was made that all complete solutions that exist for a given GSA and set of TOs lead to the same or close values of hardware expenses in the circuit of FSM. So, in the case of finding a complete solution, the algorithm finishes its work, and the found result is considered the best among all possible ones.

## 4 EXPERIMENTS

As part of experimental research, the proposed algorithm for finding partial solutions to the problem of algebraic synthesis of FSM with DT (Fig. 6) was implemented as a Python program. The implementation of the algorithm is universal and can be applied to any GSA specified by a file in the KISS format [15]. The purpose of the experiments was to reveal the ability of the algorithm to find partial solutions with the minimum possible number of uncovered transitions. The course of the experiments was as follows.

1. The FSM is given by GSA *G* (Fig. 1).

2. The program implements an enumeration of all possible variants of coding states with three-bit binary codes. In the process of enumeration, the set of TOs remained fixed. When finding a complete solution, the program continues the further enumeration of variants, and does not interrupt the search, as the algorithm predicts.

3. To determine the number of uncovered transitions, the matrix approach proposed in [10] is used. For fast processing of matrices, the standard Numpy library of the Python language is used.

4. Since GSA *G* contains only 12 transitions, the program counts not only the number of complete solutions, but also the number of partial solutions that contain one or two uncovered transitions. Such solutions for small-sized GSA can be considered as close as possible to a complete solution. The number of found complete solutions is also counted.

5. The experiment was repeated for different sets of transition operations.

Each experiment (each enumeration of all variants of states encoding for a certain set of TOs) lasted, in the case of GSA *G*, about one second. During this time, according to (5), 40320 state coding variants were enumerated.

## 5 RESULTS

Table 1 contains the results of research on the number of complete and partial solutions for GSA *G* and three TOs. Each row of the table 1 corresponds to the results of one conducted experiment. The table columns contain the following data:

1. The “Transitions operations” column shows three transitions operations used in each of the experiments. They are artificially selected so that there is at least one complete solution for them. Operations have designations similar to the designations in Fig. 2. For example, the mark “$\vee\, 110_2$” means the disjunction operation of the current state code with the binary constant 110.

2. In the “0” column the number of complete solutions found (i.e., solutions that have zero non-covered transitions) is indicated.

3. Columns “–1” and “–2” indicate the number of found partial solutions that have one or two non-covered transitions, respectively.

Table 1 – Number of complete and partial solutions for GSA *G* with using three TOs

| Transitions operations | | | 0 | –1 | –2 |
|---|---|---|---|---|---|
| + 1 | + 2 | × 2 | 1 | 18 | 80 |
| + 1 | + 2 | × 4 | 1 | 17 | 58 |
| + 1 | + 2 | + 4 | 8 | 16 | 184 |
| + 1 | + 2 | $\oplus\, 101_2$ | 8 | 56 | 200 |
| + 1 | + 2 | $\oplus\, 011_2$ | 4 | 20 | 122 |
| + 1 | + 2 | $\oplus\, 110_2$ | 4 | 4 | 58 |
| + 1 | + 2 | $\vee\, 111_2$ | 1 | 16 | 50 |
| + 1 | + 2 | $\vee\, 110_2$ | 2 | 10 | 41 |
| + 1 | + 2 | $\&\, 001_2$ | 2 | 10 | 41 |
| + 1 | + 2 | $\&\, 100_2$ | 2 | 12 | 70 |
| + 1 | + 2 | ÷ 2 | 2 | 19 | 109 |
| + 1 | + 2 | ÷ 4 | 1 | 19 | 78 |
| + 5 | $\vee\, 010_2$ | $\oplus\, 110_2$ | 4 | 4 | 58 |
| + 1 | $\vee\, 010_2$ | $\oplus\, 110_2$ | 4 | 4 | 58 |
| + 3 | + 6 | $\vee\, 010_2$ | 4 | 4 | 58 |
| + 5 | $\oplus\, 010_2$ | $\oplus\, 110_2$ | 8 | 8 | 160 |

Content of Table 1 demonstrates that in most cases the number of partial solutions with one uncovered transition exceeds the number of complete solutions. The number of partial solutions with two uncovered transitions in the conducted experiment always exceeds the number of solutions with one uncovered transition.

The fact that the values in the “–2” column exceed the values in the “–1” and “0” columns allows us to expect that during a full search, a partial solution with two uncovered transitions will be found earlier (and in less time) than solution with one uncovered transition or a complete solution. If, with two uncovered transitions, the resulting circuit of the FSM still has a satisfactory gain in hardware expenses, the search for such partial solutions can significantly speed up the process of algebraic synthesis.

Additional experiments showed that the number of solutions with three uncovered transitions is several times greater than the value in the “–2” column. Therefore, the search for less efficient solutions takes less time on average than the


DOI 10.15588/1607-3274-2024-4-14

search for more efficient solutions. This is relevant in cases of GSA of large size, for which a complete enumeration of state coding variants is impossible within acceptable time limits.

Note that the third line from the bottom of the table shows 4 found complete solutions, one of which is shown as an example in Fig. 2.

The purpose of the following experiment is to determine the number of partial solutions with one and two uncovered transitions, provided that no complete solution exists for a given set of TOs (column "0" contains zero values). The results of the experiment are shown in the Table 2, the structure of which is similar to Table 1.

Table 2 – Number of partial solutions for GSA *G* in the absence of complete solutions

| Transitions operations | | | 0 | –1 | –2 |
|---|---|---|---|---|---|
| +1 | +2 | +3 | 0 | 40 | 312 |
| + 2 | + 3 | + 4 | 0 | 24 | 184 |
| + 2 | + 3 | × 2 | 0 | 0 | 25 |
| + 1 | × 4 | ÷ 2 | 0 | 0 | 26 |
| + 1 | $\oplus 011_2$ | ÷ 2 | 0 | 1 | 11 |
| + 2 | $\vee 011_2$ | ÷ 2 | 0 | 1 | 7 |
| + 1 | $\& 011_2$ | $\oplus 101_2$ | 0 | 3 | 28 |
| + 3 | × 2 | $\oplus 110_2$ | 0 | 1 | 35 |
| $\vee 001_2$ | $\& 010_2$ | $\oplus 011_2$ | 0 | 0 | 4 |
| $\oplus 101_2$ | $\oplus 110_2$ | $\oplus 111_2$ | 0 | 0 | 96 |
| + 1 | + 5 | $\oplus 100_2$ | 0 | 128 | 320 |
| × 4 | ÷ 2 | $\oplus 110_2$ | 0 | 2 | 12 |
| × 2 | ÷ 4 | $\oplus 011_2$ | 0 | 2 | 12 |
| + 2 | ÷ 2 | $\oplus 011_2$ | 0 | 2 | 54 |
| + 3 | $\& 101_2$ | $\vee 110_2$ | 0 | 2 | 18 |
| +2 | $\oplus 011_2$ | $\vee 100_2$ | 0 | 2 | 20 |

Content of Table 2 demonstrates for each of the considered sets of operations the existence of partial solutions in the absence of complete solutions. This allows for the synthesis of FSM with DT, provided that the partial solution leads to a decrease in hardware expenses in the FSM circuit.

## 6 DISCUSSION

The proposed algorithms for finding complete and partial solutions to the problem of algebraic synthesis of FSM with DT are based on the assumption that the set of transitions operations is given and fixed. Even under this condition, the number of variants of states encoding for GSA of medium size is too large for the practical application of algorithms. Reducing the time to perform enumerations of variants is possible thanks to the application of methods and algorithms of partial enumeration, as well as due to the parallelization of the search process with an orientation to multiprocessor systems.

Meanwhile, fixing the set of TOs in the general case is not mandatory. Nothing prevents you from making an "external" enumeration of various TOs, inside which there will be an "internal" enumeration of states coding variants. This, on the one hand, will increase the total number of iterations of the algorithm, and on the other hand, it can contribute to faster finding of effective solutions.

Algorithms for the rational formation of a set of TOs can be developed instead of enumeration of TOs. In this aspect, the algorithms proposed in this article will allow, on the example of small GSAs, to investigate the effectiveness of the application of certain TOs, spreading the obtained results to large GSAs.

## CONCLUSIONS

The article proposes a solution to the scientific problem of developing algorithms for finding complete and partial solutions to the problem of algebraic synthesis of a finite state machine with datapath of transitions. The conducted research is a component of a broader scientific problem devoted to the design automatization of this class of finite state machines.

**The scientific novelty** of the article lies in the fact that for the first time algorithms for finding complete and partial solutions to the problem of algebraic synthesis have been developed, implemented and researched. Research of test GSA using the developed algorithms showed that for various sets of transitions operations, the number of partial solutions significantly exceeds the number of complete solutions. This emphasizes the relevance of the development of algorithms for the search of partial solutions using a shortened enumeration of state coding variants.

**The practical use** of obtained results is possible in the development of methods and algorithms of synthesis of FSM with DT within the framework of specialized CADs of digital control devices.

**Prospects for further research** consist in solving a range of scientific and practical problems related to the optimization of work time of proposed algorithms due to development of methods for shorted enumerations of variants of coding FSM states. This will make it possible to apply algorithms and carry out algebraic synthesis of FSM with DT for large-sized GSA, as well as automatically generate VHDL code for synthesizing FSM circuit in modern elementary bases [2, 5, 8, 9, 16].

## ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Vasyl' Stus Donetsk National University "Methods, algorithms and tools of computer-aided design of control units of computing systems" (state registration number 0122U200085).

## REFERENCES

1. Bailliul J., Samad T. Encyclopedia of Systems and Control. Springer, London, UK, 2015, 1554 p. DOI: https://doi.org/10.1007/978-1-4471-5058-9
2. Czerwinski R., Kania D. Finite state machines logic synthesis for complex programmable logic devices. Berlin, Springer, 2013, 172 p. DOI: https://doi.org/10.1007/978-3-642-36166-1
3. Baranov, S. Logic and System Design of Digital Systems. Tallin, TUTPress, 2008, 267 p.
4. Micheli G. D. Synthesis and Optimization of Digital Circuits. McGraw-Hill, Cambridge, MA, USA, 1994, 579 p.
5. Minns P., Elliot I. FSM-Based Digital Design Using Verilog HDL. JohnWiley and Sons, Hoboken, NJ, USA, 2008, 408 p. DOI: https://doi.org/ 10.1002/9780470987629
6. Skliarova I., Sklyarov V., Sudnitson A. Design of FPGA-based circuits using hierarchical finite state machines. Tallinn, TUT Press, 2012, 240 p.
7. Klimovich A. S., Solov'ev V. V. Minimization of mealy finite-state machines by internal states gluing, *Journal of Computer and Systems Science International*, 2012, Volume 51, pp. 244–255. DOI: https://doi.org/10.1134/S1064230712010091


DOI 10.15588/1607-3274-2024-4-14

8. Grout, I. Digital Systems Design with FPGAs and CPLDs / I. Grout. – Elsevier Science: Amsterdam, The Netherlands, 2008. – 784 p. DOI: https://doi.org/10.1016/B978-0-7506-8397-5.X0001-3
9. Kubica M., Opara A. and Kania D. Technology Mapping for LUT-Based FPGA, *Lecture Notes in Electrical Engineering*, Springer, Cham, 2021, Vol. 713, 207 p. DOI: https://doi.org/10.1007/978-3-030-60488-2
10. Baranov S. Logic Synthesis for Control Automata. Dordrecht, Kluwer Academic Publishers, 1994, 312 p.
11. Barkalov A. A., Babakov R. M. Operational formation of state codes in microprogram automata, *Cybernetics and Systems Analysis*, 2011, Volume 47 (2), pp. 193–197. DOI: https://doi.org/10.1007/s10559-011-9301-y
12. Barkalov A. A., Titarenko L. A., Babakov R. M. Synthesis of Finite State Machine with Datapath of Transitions According to the Operational Table of Transitions, *Radio Electronics, Computer Science, Control*, 2022, No. 3 (62), pp. 109–119. DOI: https://doi.org/10.15588/1607-3274-2022-3-11
13. Barkalov A. A., Babakov R. M. A Matrix Method for Detecting Formal Solutions to the Problem of Algebraic Synthesis of a Finite-State Machine with a Datapath of Transitions, *Cybernetics and Systems Analysis*, 2023, Volume 59 (2), pp. 190–198. DOI: https://doi.org/10.1007/s10559-023-00554-6
14. Barkalov A. A., Titarenko L. A., Babakov R. M. Synthesis of VHDL-Model of a Finite State Machine with Datapath of Transitions, *Radio Electronics, Computer Science, Control*, 2023, No. 4 (67), pp. 135–147. DOI: https://doi.org/10.15588/1607-3274-2023-4-13
15. Yang S. Logic synthesis and optimization benchmarks User Guide Version 3.0. Tech. rep. Microelectronics Center of North Carolina, P.O. Box 12889, Research Triangle Park, NC 27709, 1991.
16. Rawski M., Selvaraj H., Luba T. An application of functional decomposition in ROM-based FSM implementation in FPGA devices, *Journal of System Architecture*, 2005, Volume 51, pp. 424–434. DOI: https://doi.org/10.1016/j.sysarc.2004.07.004

Received 05.09.2024.
Accepted 07.11.2024.

УДК 004.94 : 004.2

## АЛГОРИТМІЧНІ ВІДМІННОСТІ ПОВНОГО І ЧАСТКОВОГО АЛГЕБРАЇЧНОГО СИНТЕЗУ МІКРОПРОГРАМНОГО АВТОМАТА З ОПЕРАЦІЙНИМ АВТОМАТОМ ПЕРЕХОДІВ

**Бабаков Р. М.** – д-р техн. наук, доцент, доцент кафедри інформаційних технологій Донецького національного університету імені Василя Стуса, м. Вінниця, Україна.

**Баркалов О. О.** – д-р техн. наук, професор, професор Інституту комп'ютерних наук та електроніки університету Зеленогурського, м. Зельона Гура, Польща.

**Тітаренко Л. О.** – д-р техн. наук, професор, професор Інституту комп'ютерних наук та електроніки університету Зеленогурського, м. Зельона Гура, Польща.

**Войтенко М.О.** – студент факультету інформаційних і прикладних технологій Донецького національного університету імені Василя Стуса, м. Вінниця, Україна.

### АНОТАЦІЯ

**Актуальність.** Розглянуто задачу алгоритмізації пошуку формальних розв'язків задачі алгебраїчного синтезу мікропрограмного автомата з операційним автоматом переходів. Запропоновано поняття повного та часткового розв'язків цієї задачі. Об'єктом дослідження є автоматизований синтез автомата в частині функції переходів без урахування функції виходів. В основі алгебраїчної реалізації функції переходів знаходиться авторський підхід до перетворення кодів станів за допомогою множини арифметико-логічних операцій. Пошук формальних розв'язків задачі алгебраїчного синтезу є складним процесом, що потребує використання відповідних методів і алгоритмів, спрямованих на спеціальне кодування станів та зіставлення операцій переходів окремим автоматним переходам. Використання часткових розв'язків задачі алгебраїчного синтезу може сприяти зменшенню часу роботи таких алгоритмів та зменшенню загального часу проектування цифрових пристроїв керування на базі мікропрограмного автомата з операційним автоматом переходів.

**Мета.** Розробка і дослідження алгоритмів пошуку повного і часткового розв'язків задачі алгебраїчного синтезу мікропрограмного автомата з операційним автоматом переходів.

**Метод.** В основу дослідження покладено структуру мікропрограмного автомата з операційним автоматом переходів. Синтез схеми автомата передбачає попереднє розв'язання задачі алгебраїчного синтезу. Результатом є так званий формальний розв'язок цієї задачі, який містить в собі дві складових. Першою складовою є визначені коди станів, другою складовою – арифметико-логічні операції, зіставлені окремим автоматним переходам. Автомат може бути синтезований в тому випадку, якщо при заданих кодах станів їх перетворення в процесі виконання переходів можливе за допомогою заданої множини операцій. Перевірка цієї можливості здійснюється за допомогою відомого матричного підходу. Від передбачає формування і поелементне зіставлення двох матриць – матриці переходів і об'єднаної матриці операцій. В результаті формується матриця покриття, яка відображає можливість реалізації (покриття) автоматних переходів за допомогою заданих арифметико-логічних операцій. Зміна способу кодування станів або набір операцій може давати інші розв'язки задачі алгебраїчного синтезу з більшою чи меншою кількістю покритих автоматних переходів.

**Результати.** На прикладі абстрактної граф-схеми алгоритму керування продемонстровано, що розв'язком задачі алгебраїчного синтезу мікропрограмного автомата з операційним автоматом переходів може вважатись ситуація, коли один або більше автоматних переходів не можуть бути реалізовані за допомогою заданого набору операцій. Таку ситуацію запропоновано називати частковим розв'язком задачі алгебраїчного синтезу. Реалізація усіх без винятку переходів за допомогою заданих операцій дає повний розв'зок цієї задачі, однак кількість повних розв'язків завжди є значно меншою за кількість часткових розв'язків. Отже, в загальному випадку пошук повних розв'язків займає набагато більше часу і до того ж є не завжди можливим.

**Висновки.** Проєктування логічної схеми мікропрограмного автомата з операційним автоматом переходів можливе у випадку наявності повного або часткового розв'язку задачі алгебраїчного синтезу. У випадку часткового розв'язку ті переходи, які не можуть бути реалізовані жодною з наявних операцій, реалізуються в канонічний спосіб за допомогою системи булевих рівнянь. Пошук повних розв'язків в загальному випадку потребує більше часу, ніж пошук часткових розв'язків. Це робить актуальним розробку алгоритмів і методів синтезу даного класу автоматів, основаних на пошуку часткових розв'язків задачі алгебраїчного синтезу.

**КЛЮЧОВІ СЛОВА:** мікропрограмний автомат, операційний автомат переходів, граф-схема алгоритму, арифметико-логічні операції, алгебраїчний синтез, частковий розв'язок.


DOI 10.15588/1607-3274-2024-4-14

**ЛІТЕРАТУРА**

1. Bailliul J. Encyclopedia of Systems and Control / J. Bailliul, T. Samad. – Springer: London, UK, 2015. – 1554 p. DOI: https://doi.org/10.1007/978-1-4471-5058-9
2. Czerwinski R. Finite state machines logic synthesis for complex programmable logic devices / R. Czerwinski, D. Kania. – Berlin: Springer, 2013. – 172 p. DOI: https://doi.org/10.1007/978-3-642-36166-1
3. Baranov S. Logic and System Design of Digital Systems / S. Baranov. – Tallin : TUTPress, 2008. – 267 p.
4. Micheli G. D. Synthesis and Optimization of Digital Circuits / G. D. Micheli. – McGraw-Hill : Cambridge, MA, USA, 1994. – 579 p.
5. Minns P. FSM-Based Digital Design Using Verilog HDL / P. Minns, I. Elliot. – JohnWiley and Sons : Hoboken, NJ, USA, 2008. – 408 p. DOI: https://doi.org/ 10.1002/9780470987629
6. Skliarova I. Design of FPGA-based circuits using hierarchical finite state machines / I. Skliarova, V. Sklyarov, A. Sudnitson. – Tallinn : TUT Press, 2012. 240 p.
7. Klimovich A. S. Minimization of mealy finite-state machines by internal states gluing / A. S. Klimovich, V. V. Solov'ev // Journal of Computer and Systems Science International. – 2012. – Volume 51. – P. 244–255. DOI: https://doi.org/10.1134/S1064230712010091
8. Grout I. Digital Systems Design with FPGAs and CPLDs / I. Grout. – Elsevier Science: Amsterdam, The Netherlands, 2008. – 784 p. DOI: https://doi.org/10.1016/B978-0-7506-8397-5.X0001-3
9. Kubica M. Technology Mapping for LUT-Based FPGA / M. Kubica, A. Opara and D. Kania // Lecture Notes in Electrical Engineering. – Springer, Cham. – 2021. – Vol. 713. – 207 p. DOI: https://doi.org/10.1007/978-3-030-60488-2
10. Baranov S. Logic Synthesis for Control Automata / S. Baranov. – Dordrecht : Kluwer Academic Publishers, 1994. – 312 p.
11. Barkalov A. A. Operational formation of state codes in microprogram automata / A. A. Barkalov, R. M. Babakov // Cybernetics and Systems Analysis. – 2011. – Volume 47 (2). – P. 193–197. DOI: https://doi.org/10.1007/s10559-011-9301-y
12. Barkalov A. A. Synthesis of Finite State Machine with Datapath of Transitions According to the Operational Table of Transitions / A. A. Barkalov, L. A. Titarenko, R. M. Babakov // Radio Electronics, Computer Science, Control. – 2022. – No. 3 (62). – P. 109–119. DOI: https://doi.org/10.15588/1607-3274-2022-3-11
13. Barkalov A. A. A Matrix Method for Detecting Formal Solutions to the Problem of Algebraic Synthesis of a Finite-State Machine with a Datapath of Transitions / A. A. Barkalov, R. M. Babakov // Cybernetics and Systems Analysis. – 2023. – Volume 59 (2). – P. 190–198. DOI: https://doi.org/10.1007/s10559-023-00554-6
14. Barkalov A. A. Synthesis of VHDL-Model of a Finite State Machine with Datapath of Transitions / A. A. Barkalov, L. A. Titarenko, R. M. Babakov // Radio Electronics, Computer Science, Control. – 2023. – No. 4 (67). – P. 135–147. DOI: https://doi.org/10.15588/1607-3274-2023-4-13
15. Yang S. Logic synthesis and optimization benchmarks / Yang S. // User Guide Version 3.0. – Tech. rep. Microelectronics Center of North Carolina, P.O. Box 12889, Research Triangle Park, NC 27709, 1991.
16. Rawski M. An application of functional decomposition in ROM-based FSM implementation in FPGA devices / M. Rawski, H. Selvaraj, T. Luba // Journal of System Architecture. – 2005. – Volume 51. – P. 424– 434. DOI: https://doi.org/10.1016/j.sysarc.2004.07.004


DOI 10.15588/1607-3274-2024-4-14

UDC 528.29

# METHODOLOGY FOR OPTIMIZING THE FUNCTIONING OF THE OPTOELECTRONIC SURVEILLANCE SYSTEM

**Borovyk D. O.** – Master of Computer Science, Khmelnytskyi, Ukraine.

**Borovyk O. V.** – Dr. Sc., Professor, Deputy Head of the Department of Organization of Educational and Scientific Activities of the Professional Training Department of the Administration of the State Border Guard Service of Ukraine, Kyiv, Ukraine.

**Rachok R. V.** – Dr. Sc. Professor, Professor of the Department of Communication and Information Systems, Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine.

**Basaraba I. O.** – PhD, Instructor of Foreign Languages Department, Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine.

## ABSTRACT

**Context.** Radar, thermal imaging, and video surveillance means are actively used in the protection of the state border. Together with the appropriate communication equipment, they allow to create optoelectronic surveillance systems, which are the basis for the intellectualization of border protection. The effectiveness of such systems is significantly affected by the peculiarities of their functional and structural design. A rational structural design, even in difficult physical and geographical conditions, allows for a high level of surveillance efficiency. However, the functional component also has a significant impact on improving the system efficiency. In many cases, the functioning of the elements of the optoelectronic surveillance system occurs under conditions of power supply restrictions. Such limitations determine the need for a rational choice of modes of use of certain types of surveillance equipment at certain time intervals in order to ensure effective surveillance, taking into account the time of day and weather conditions. The imperfection of the scientific and methodological apparatus for optimizing the functioning of optoelectronic surveillance systems determines the relevance of this study.

**Objective.** The aim of the work is to develop a methodology for optimizing the optoelectronic surveillance system functioning by rationally selecting the modes of operation of different types of surveillance equipment in certain time intervals, taking into account the time of day and weather conditions in which they are used.

**Methods.** The paper sets and investigates the two-criteria problem of choosing the modes of operation of different types of observation equipment of an optoelectronic surveillance system at separate time intervals, taking into account the time of day and weather conditions in which they are used, which ensures maximizing the efficiency of the optoelectronic surveillance system while minimizing the power consumed by active types of surveillance equipment in the presence of boundary restrictions on the efficiency and power consumed by the system.

The proposed indicator for assessing the effectiveness of the system allows us to assess the level of impossibility of uncontrolled crossing of the perimeter of the protected area by an intruder. The peculiarity of this methodology is the possibility of ensuring a significant reduction in the level of energy consumption by the system components due to a slight decrease in the efficiency of monitoring.

**Results.** The paper proposes an alternative approach to assessing the effectiveness of the optoelectronic surveillance system, the idea of which is that instead of assessing the effectiveness of surveillance over the entire sector of the controlled area of the border, the effectiveness of control is assessed only along the perimeter of this area. This approach significantly reduces the computational complexity of the problem of finding the value of efficiency which further simplifies the solution of problems of structural optimization of surveillance systems. A software and algorithmic implementation of the methodology for optimizing the functioning of an optoelectronic surveillance system is proposed. Using the developed software, a rational choice of modes of operation of certain types of surveillance equipment at certain time intervals was carried out taking into account the time of day and weather conditions.

**Conclusions.** The use of the proposed methodology makes it possible to optimize the modes of operation of the optoelectronic surveillance system, taking into account the limiting factors in terms of efficiency and power consumption when using the same types of surveillance equipment on all towers of the system. A possible direction for improving the methodology is its adaptation to cases where different types of surveillance equipment are used on different towers of the system.

**KEYWORDS:** optoelectronic surveillance system, methodology, optimization, efficiency, power consumption, algorithm.

## ABBREVIATIONS

SBGSU – State Border Guard Service of Ukraine;

OSS – optoelectronic surveillance system.

## NOMENCLATURE

$Y$ – set of weather conditions;

$y_i$ ($i$=[1,$u$]) – elements of the set of weather conditions;

$T_o$ – time of day;

$T_c$,$T_m$ – light, dark periods of the day;

$T$ – period of time during which the functioning of the OSS is studied;

$t_i$ ($i$=[0,$s$]) – start times of the studied elementary discrete time ranges in the time interval during which the functioning of the OSS is studied;

$t_0$ – moment of the beginning of the first elementary discrete time range in the interval $T$ (the moment of the beginning of the period $T$);

$t_s$ – start time of the last elementary discrete time range in the interval $T$;


DOI 10.15588/1607-3274-2024-4-15

$n_{j,ti}$ ($j$=[1,$k$]) – number of surveillance devices of the *j-th* type operating in the OSS during an elementary discrete time range starting at $t_i$ time;

$P_{j,ti}$ ($j$=[1,$k$]) – power consumption of electricity by the $j$-th type of observation means during an elementary discrete time range starting at $t_i$ time, characterized by specific $y_i$ weather conditions (for example, from a weather forecast service) and the period of day;

$P_T$ – power of electricity consumed by active types of OSS monitoring devices during the $T$ studied period $T$;

$P_0$ – maximum permissible power limit of the consumed electricity by active types of OSS monitoring equipment during the $T$ studied period;

$E_{j,t_i}$ – effectiveness of the $j$-th type of observation means during an elementary discrete time range starting at $t_i$ time, characterized by specific weather conditions $y_i$ (for example, from a weather forecast service) and the period of day;

$E_T$ – effectiveness of the OSS during the $T$ studied period in detecting the fact of violation of the border section perimeter;

$E_0$ – minimum permissible limit of OSS efficiency during the $T$ studied period;

$pv_k(p_{ij}, y_i, x, y)$ – probability of detecting a target by a surveillance device of the $k$-th type located at $p_{ij}$ point, in an elementary section of the studied control area with a centre at point ($x$, $y$) and under weather conditions $y_i$.

## INTRODUCTION

Reliable protection of Ukrainian borders is an absolute guarantee of their inviolability in peacetime. In today's environment, effective protection is impossible without proper engineering equipment and border arrangements. This equipment, in particular, includes the latest technical surveillance equipment, such as thermal imaging cameras, modern radar stations and video cameras, as well as modern communication means. The effectiveness of a surveillance system built with such equipment depends heavily on the perfection of its structural and functional organization, as well as the compliance of the chosen mode of use of the equipment with the weather and other conditions in which the surveillance is carried out. When solving the problems of structural optimization of the surveillance system, the use of the effectiveness indicator, which is focused on assessing the functioning of the entire sector of the border area under study is not always rational. In most cases for the effective organization of border protection, the urgent task is to ensure the most effective monitoring of the perimeter of the sector of the studied area, and in case of its violation, the sector of the entire area limited by the specified perimeter. Therefore, it is important to clarify an objective indicator of the effectiveness of monitoring the perimeter of the border area.

In some cases, the operation of OSS must be carried out with autonomous power supply of equipment through the use of renewable energy sources. All of this makes it important to optimize the use of a complex of surveillance equipment while maximizing their efficiency and minimizing their energy consumption.

It is assumed that under fixed observation conditions, one of the monitoring devices can provide an efficiency close to the total efficiency provided by the simultaneous operation of all types of devices. In this case, a significant improvement in the energy efficiency of the system can be achieved by slightly reducing the efficiency of monitoring when operating only one of the most efficient means of monitoring in specific conditions.

**The object of research** is the functioning of the optoelectronic surveillance system.

**The subject of research** are the methods of optimizing the functioning of the optoelectronic surveillance system.

**The purpose of this work** is to develop a methodology for optimizing the functioning of the optoelectronic surveillance system by rationally selecting the modes of operation of different types of surveillance equipment at certain time intervals taking into account the time of day and weather conditions in which they are used. This methodology provides the establishment of an indicator of the effectiveness of the OSS in detecting an intruder at the border of the area where control is to be ensured.

## 1 PROBLEM STATEMENT

Suppose that OSS consists of $n$ observation towers, the location of which is stationary and optimized for the problem under study [1]. Each tower has $k$ types of surveillance equipment. There are quite a few possible types of surveillance equipment. Typically, optical, optoelectronic, thermal imaging and radar stations are used at OSS. The effectiveness of surveillance equipment varies depending on the conditions of use. The conditions of use include weather conditions and time of day.

Depending on the time of year, the duration of the light $T_c$ and dark $T_m$ periods of the day $T_\partial$ are different and variable.

Certain types of surveillance equipment on the towers can operate independently of each other and regardless of what types of equipment are operating on other towers. For the purposes of this study, we will assume that during an elementary discrete time range starting at $t_i$ time, the same type of surveillance equipment can operate on all towers.

Given that the $E_T$ indicator is for the $T$ period, which can be either partially occurred or fully occurred in the future, to find $E_T$, you can use the values $y_i$, which are determined from the weather forecast service for each elementary discrete time range starting at $t_i$ time.

The task is to determine for $T$ period, taking into account the set of $y_i$ values, which are determined from the weather forecast service and are characteristic of $T$ period, the types of observation devices on the OSS towers that will provide


DOI 10.15588/1607-3274-2024-4-15

$$E_T \to \max, \quad (1)$$

$$P_T \to \min. \quad (2)$$

with restrictions

$$E_T \geq E_0, \quad (3)$$

$$P_T \leq P_0. \quad (4)$$

In the optimization problem (1)–(4)

$$E_T = \sum_{\substack{t_s \\ t_i=t_0, \\ j}} n_{j,t_i} E_{j,t_i}, \quad (5)$$

$$P_T = \sum_{\substack{t_s \\ t_i=t_0, \\ j}} n_{j,t_i} P_{j,t_i}. \quad (6)$$

In the right-hand sides of expressions (5), (6), the summation is performed by the values that determine the active surveillance means for each elementary discrete time range starting at $t_i$ time.

The solution of problems (1)–(6) makes it possible to plan the operating modes of the OSS monitoring elements for the $T$ period, as well as (if necessary) to adjust these modes if the time $T$ interval is sufficiently long, or the data from the weather forecast service is not accurate enough, or is accurate enough only for small periods of time.

## 2 LITERATURE REVIEW

A significant number of scientific papers are devoted to the study of improving the functioning of surveillance systems. A general overview of surveillance systems is given in [2]. This study examines the features of the construction of surveillance systems and the basic procedures for processing information in them. Also, [2] identifies trends in the development of surveillance systems and relevant areas for further research in this area.

Paper [3] discusses the improvement of air target signal detection in complex surveillance systems.

Many researchers focus on the issues of efficient information processing in surveillance systems to automate object detection. For example, work [4] considers the use of neural networks with deep-learning algorithms for this purpose. Similarly to the previous study, the issues of image classification using neural networks are reflected in [5]. An interesting and promising area of development of modern surveillance systems is the acquisition and processing of images from unmanned aerial vehicles [6]. Methods of image comparison with mask construction are discussed in [7].

In the construction of surveillance systems, especially when using surveillance equipment with limited communication capabilities, attention is paid to optimizing the transmission of video streams. Thus, in [8], a model for classifying informative image segments is proposed, on the basis of which information redundancies in aerial photographs are eliminated.

Some studies in the field of surveillance systems [9–10] are devoted to the integration of images in different parts of the spectral range. Thus, [9] proposes to improve the methods of forming integrated multispectral images by taking into account their local features. Study [10] identifies possible ways to increase the information content of images based on the determination of the operating modes of multispectral optical channels.

A number of works [1, 11–12] address the issues of structural and functional optimization of OSS used in border protection. However, these studies use an efficiency indicator that reflects the ability to conduct surveillance over the entire area to be controlled. At the same time, in many partial variants of the use of surveillance systems, it is sufficient to reliably control only the perimeter of the site. And this requires an appropriate adjustment of the performance indicator.

Along with the structural and functional optimization of surveillance systems [12], it is also advisable to rationally choose the mode of operation of surveillance means in accordance with weather and other conditions in order to reduce energy consumption and ensure sufficiently high efficiency. All this determines the relevance of further relevant research.

## 3 MATERIALS AND METHODS

An important component of the OSS, which significantly affects the overall efficiency of its use, is a complex of towers with installed surveillance equipment, the functioning of which is based on the use of electromagnetic radiation in different frequency range (Figure 1).

Depending on the conditions in which the surveillance is conducted, the surveillance equipment of the OSS provides different efficiency. Each of the devices consumes a certain amount of power from the power source. In some cases, when the power supply is autonomous, its resource is limited. This prompts the search for ways to optimize power consumption.

Under certain weather conditions, the effectiveness of surveillance using a particular type of sensor can be significantly better than that provided by other types of sensors. In this case, from the point of view of minimizing energy consumption, it is inappropriate to use all available types of surveillance equipment at the same time. It is assumed that due to the rational choice of the type of observation means for each elementary discrete time range starting at $t_i$ time, corresponding to the weather and other conditions in which the observation is carried out, and possibly a slight decrease in efficiency, significant savings in electricity consumption can be


DOI 10.15588/1607-3274-2024-4-15

achieved. Such savings can be extremely important in cases of using renewable energy sources (sun, wind, etc.).

For example, if three types of surveillance equipment are used on OSS towers, the option of using separate types of equipment on separate time ranges without taking into account the restrictions on the power consumption by them can be determined as follows. First of all, the effectiveness of the observation means of $j=1$, $j=2$, $j=3$ type on elementary discrete time points starting at $t_i$ time moments during the studied period $T$ is determined (see Fig. 2), and then the option of using observation means (see Table 1) is established as a choice of the type of means $j=1$, $j=2$, $j=3$, for each elementary discrete time band starting at $t_i$ time points which is characterized by maximum efficiency.

Figure 1 – Example of OSS tower and devices

It should be noted that if it is necessary to take into account restrictions on electricity consumption, the option of using surveillance equipment may differ from that given in Table 1. For example, if the power consumption of the $j=2$ device is significantly higher than the power consumption for the $j=3$ device , then the option of using surveillance equipment may be as given in Table 2.

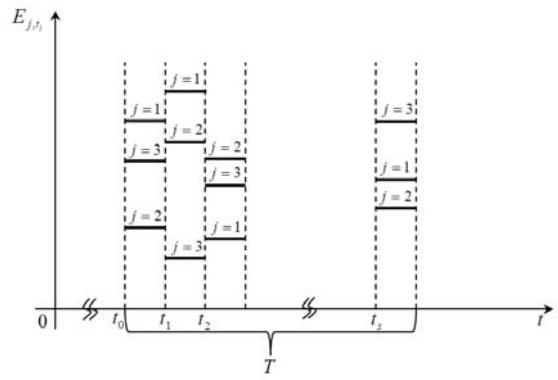



Figure 2 – A possible variant of realization of efficiency by means of observations of $j=1$, $j=2$, $j=3$ type on elementary discrete time ranges with a start at $t_i$ time during the studied period $T$

Table 1

| The research period | $T$ | | | | |
|---|---|---|---|---|---|
| The beginning of the studied elementary discrete time ranges | $t_0$ | $t_1$ | $t_2$ | … | $t_s$ |
| The $j$ type of observation device that is appropriate for use in the elementary discrete time range under study | 1 | 1 | 2 | … | 3 |

Table 2

| The research period | $T$ | | | | |
|---|---|---|---|---|---|
| The beginning of the studied elementary discrete time ranges | $t_0$ | $t_1$ | $t_2$ | … | $t_s$ |
| The $j$ type of observation device that is appropriate for use in the elementary discrete time range under study | 1 | 1 | 3 | … | 3 |

At the same time, the decision to use a particular type of surveillance device is formed on the basis of solving the optimization problem (1)–(6). That is why the task set out in this paper becomes relevant.

The analysis of expressions (5), (6) allows us to conclude that the solution of the optimization problem (1)–(6) directly depends on the type of $E_{j,t_i}$ function .

The task of determination this type of function was studied in [1, 11–12]. In these works, an indicator similar to $E_{j,t_i}$, was determined on the basis of a probabilistic approach, according to which, for each observation means, the dependence of the probability of detecting an object on the distance to it was experimentally determinated, in accordance with the observation conditions. This approach made it possible to take into account both deterministic and stochastic factors that affect the detection of objects in specific observation conditions. In these works, such an indicator had the following form:

$$E_S = \frac{\sum\limits_{(x,y)\in S_m} \left(1-\prod\limits_{i,j}(1-pv_k(p_{ij},y_t,x,y))\right)}{nSm}. \qquad (7)$$


DOI 10.15588/1607-3274-2024-4-15

This efficiency indicator provided the possibility of averaging the probability of detecting objects over all $(x,y)$ discrete elements of $S_m$ observation band area (Figure 3).

The $E_s$ calculation takes into account the possible fertilizing effect of the terrain. This is ensured by using the capabilities of modern information technologies.

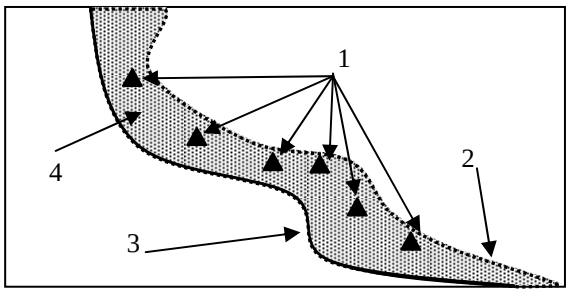



Figure 3 – Diagram of the layout of OSS elements:
1 – observation towers, 2 – perimeter of observation,
3 – border line, 4 – observation lane

If the element of $(x,y)$ location is not visible (i.e., if the $(x,y)$ element belongs to the invisible zone), the corresponding value of $pv_k(p_{ij}, y_i, x, y)$ is 0. If this element is visible, then the following formula is used to calculate $pv_k(p_{ij}, y_i, x, y)$

$$pv_k(p_{ij}, y_t, x, y) = \begin{cases} P_1(R(p_{i,j}, x, y), k, y_i), (x, y) \notin T_{p_{ij}}, \\ 0, (x, y) \in T_{p_{ij}}, \end{cases} \quad (8)$$

where

$$P_1(R, k, y_i) = 1 - \frac{1}{\sigma_R(k, y_i)\sqrt{2\pi}} \int_{R_{\min}}^{R} e^{-\frac{(R - m_R(k, y_i))^2}{2\sigma_R^2(k, y_i)}} dR. \quad (9)$$

In expression (7), $nSm$ this is the number of discrete elements $(x,y)$ of $S_m$.observation band area.

Expression (8) implements a probabilistic description of the effectiveness of detecting an object by the $k$-th observation means depending on the distance $R$ to it. The $R$ distance is determined by the coordinates of the $p_{ij}$ observation point (corresponding to the OSS tower where the equipment is installed) and the coordinates $(x,y)$ of the center of the elementary area under observation According to the observation conditions and the type of observation means, the parameters of the distribution law (the mean and variance of the random variable $R$) are selected in (9). These parameters are set experimentally on the basis of statistical processing of tests of the corresponding observation means in different conditions. When applying formula (8), geoprocessing is used, which consists in determining the correspondence of the area $(x,y)$ to the set of invisibility zones from the tower located at $p_{ij}$ point.

However, such a determination of the efficiency indicator (7) requires finding the visibility of each of its discrete elements $(x,y)$. This task for $S_m$ large areas even when using optimized algorithms for determining visibility has a great computational complexity.

Taking into account the fact that in many cases it is not necessary to observe the whole area of the observation strip when controlling the border section, and to detect objects it is enough to observe only the perimeter of the section, then taking into account the probabilistic approach implemented in dependencies (7)–(9), the efficiency indicator can be represented as

$$E_{j,t_i} = \frac{\sum_{(x,y)\in l_m} (1 - \prod_{i,j} (1 - pv_k(p_{ij}, y_i, x, y)))}{nl_m}. \quad (10)$$

In contrast to (7), in (10), the summation and calculation of $pv_k(p_{ij}, y_i, x, y)$ with the determination of visibility $(x, y)$ is not performed over the entire area of the $S_m$ observation sector, but only along $l_m$ perimeter.

Thus, this approach defines an original methodology for optimizing the functioning of the optoelectronic surveillance system by rationally selecting the modes of operation of different types of surveillance equipment at certain time intervals taking into account the time of day and weather conditions in which they are used.

In the above methodology $E_{j,t_i}$ the quantities may have a complex characteristic (not necessarily the one shown in Fig. 2). In this case the application of the approach described in Fig. 2 and Tables 2–3 is not obvious. Consequently, solving problems (1)–(6) can be significantly complicated even though the computational complexity of the problem is reduced by conducting a study on the border of the perimeter of the observation sector $S_m$.

In order to specify the proposed methodology, we will conduct additional analysis.

The time of day has a significant impact on the effectiveness of surveillance equipment. For example, if we compare two types of surveillance cameras operating in the visible and far infrared ranges (thermal imaging cameras), then during daylight hours, an optical range camera will be preferable for detecting objects, as it provides significantly higher resolution. At the same time, the simultaneous use of these cameras will lead to unnecessary energy consumption. However, the situation will change dramatically with the onset of darkness. This can be estimated from the graph shown in Fig. 4.

During daylight hours, the probability of detecting an object using an optical range camera under favorable weather conditions is quite high (Figure 4 – solid line). However, this probability decreases significantly when the light level decreases. For a thermal imaging camera (Figure 4 – dotted line), the situation is significantly different. At night, a thermal imager can clearly distinguish objects even in the dark, making it an indispensable tool for nighttime surveillance. However, during the day, when the environment is warm, the thermal imager may have more difficulty in detecting objects because the temperature difference between the objects and the environment is smaller.


DOI 10.15588/1607-3274-2024-4-15

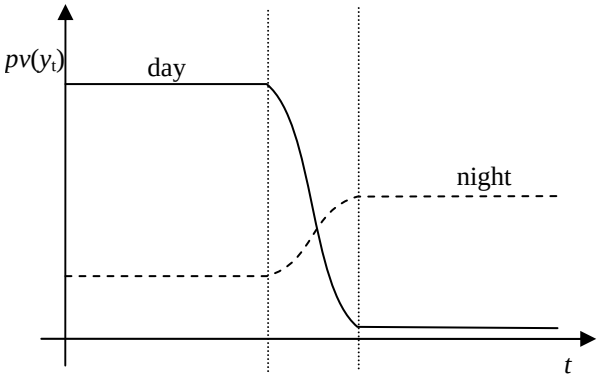



Figure 4 – Dependence of detection probability at a fixed distance for two different types of surveillance equipment on time

Weather conditions *Y* also have a significant impact on the effectiveness of surveillance equipment. It should be noted that the dependence of efficiency on time for different types of surveillance equipment, taking into account possible changes in the conditions of observation, can be complex (Figure 5). This is influenced by a combination of different factors.

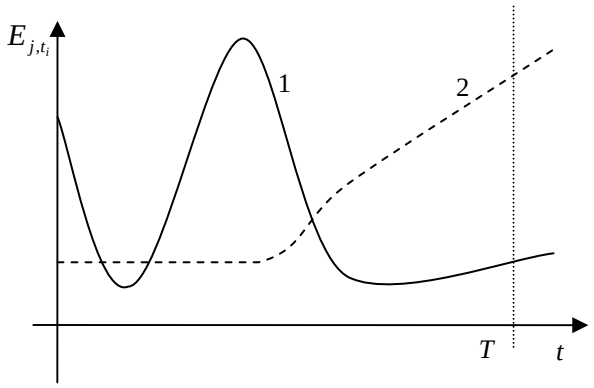



Figure 5 – Possible dependencies of the efficiency indicator (10) for two different types of surveillance equipment on time taking into account the conditions of observation

During the studied *T* time interval, within which one of the types of surveillance means is chosen, the efficiency for different surveillance means can change dramatically. Thus, at the beginning of the studied interval (Figure 5 – solid line), the best efficiency is provided by the 1-st surveillance means. However, this effectiveness for the 1-st means decreases sharply over time, and the use of the 2-nd means becomes preferable. However, after a short period of time, the 1-st means again provides greater efficiency. All of this makes it difficult to choose the surveillance tools for individual discrete time periods.For a rational choice of an observation tool, one can use a comparison of the mathematical expectations of the $E_{j,t_i}$ value over the *T* time interval. However, it should be noted that the minimum and maximum efficiency are also important. Given that the minimum, maximum, and average efficiencies for individual instruments are not always such that they unambiguously determine the choice of an instrument, it is necessary to determine a selection criterion.

The analysis allows us to assume that it is advisable to use the mathematical apparatus of fuzzy values for the rational choice of an active surveillance means. The effectiveness of each type of surveillance means in the studied interval can be determined by the $E_T$,fuzzy value which is described by a triangular function that takes into account the minimum, average and maximum values of the corresponding effectiveness. In this case, the membership function for fuzzy efficiency will be given as

$$\mu(E_T) = \begin{cases} 0, E_T < \min E_{j,t_i}, \\ 1, E_T = M[E_{j,t_i}], \\ 0, E_T > \max E_{j,t_i}, \end{cases} \quad (11)$$

where M[] is the operator of mathematical expectation.

The concept of fuzzy maximum is used to compare the fuzzy efficiencies (11) of different types of surveillance means.

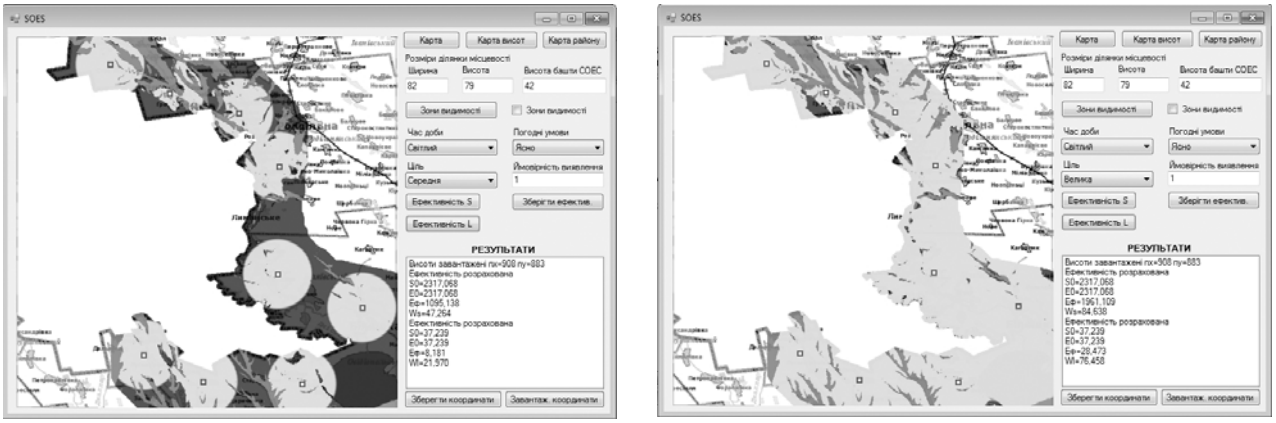
Thus, the concretization of the proposed methodology for optimizing the functioning of the optoelectronic surveillance system through the rational choice of the mode of using different types of surveillance equipment in accordance with weather and other conditions concerns the mechanisms for determining $E_{j,t_i}$ values.

## 4 EXPERIMENTS

In order to automate the calculation of the value (10), a software application was developed (Figure 6).

The application implements geoprocessing to determine the areas of the sectors visible from the observation towers. Observation conditions including weather conditions are set using the appropriate interface elements and, accordingly, the program selects the parameters of the probabilistic description of the operation of the observation means. In addition to calculating a predefined indicator of the effectiveness of monitoring the area (7), the application implements the calculation of the effectiveness of controlling the perimeter of a certain area (10). For convenience, the storage and reading of the locations of observation towers, which can be edited, is implemented.


DOI 10.15588/1607-3274-2024-4-15



a b

Figure 6 – An application for automating efficiency finding (10)

## 5 RESULTS

Figure 6 shows some examples of the results of calculating the indicators of observation efficiency by area (7) and by perimeter (10).

When calculating the corresponding indicators, the window displays the total area of the areas for which $pv_k(p_{ij}, y_i, x, y)$ is calculated. These calculations constitute the main computational complexity in solving the problem. From the above examples of application use, it can be concluded that the computational complexity is reduced by more than 60 times.

An interesting fact is that in all variants, the absolute values of the efficiency indicator (10) are lower than the value of (7). On the one hand, this is explained by the fact that the task of structural optimization of the OSS – the placement of observation towers was carried out with the maximization of (7). However, the main reason is that the physical meaning of indicators (7) and (10) is to average the values of the probability of detecting an object over all elementary areas that make up the study area (in the case of (7)) or perimeter (for indicator (10)). At the same time, a significant part of the area is located relatively close to the observation towers, where a high probability of detecting targets is ensured. Therefore, the average value can be quite close to 1. In the case of performance indicator (10), most areas of the perimeter of the surveillance strip are located at a considerable distance from the towers, and therefore the averaged values are significantly lower.

However, within the framework of using each of the efficiency indicators, it is possible to compare different means of surveillance in different conditions and choose the most appropriate mode of operation of the system in accordance with these conditions.

To choose such a mode, a technique is proposed that, using the mathematical apparatus of fuzzy values, ensures the determination of the observation device that provides the best observation efficiency over $T$ time period.

## 6 DISCUSSION

The methodology for optimizing the functioning of the optoelectronic surveillance system developed in this paper is unique due to the use of two optimization criteria, unlike the existing ones [1, 11–12]. The solution of a two-criteria optimization problem is not only of practical importance, but also of theoretical significance which consists in the development of the theory of computational processes. In the studied problem, a new indicator for assessing the effectiveness of the system's functioning is also used for the first time, which, unlike the existing ones allows assessing the level of impossibility of uncontrolled crossing of the perimeter of the protected area by an intruder.

The peculiarity of the studied problem, which somewhat idealizes the process of using the monitoring elements of OSS, is that the same type of monitoring equipment was chosen at all observation towers. The issue of establishing the dependence of observation effectiveness on the choice of different means on different towers under certain observation conditions (a generalized problem) is currently unexplored. However, the proposed approach to solving the problem studied in this paper can be applied as one of the stages of solving the generalized problem.

## CONCLUSIONS

The problem of optimizing the functioning of the optoelectronic surveillance system in accordance with its formulation in this study can be considered solved.

**The scientific novelty** of the obtained results lies in a new approach to determining the efficiency of detecting objects in the observation sector by an optoelectronic surveillance system.

The paper sets and investigates a two-criteria problem of choosing the modes of operation of different types of surveillance equipment of an optoelectronic surveillance system at certain time intervals, taking into account the time of day and weather conditions in which they are used, which ensures maximizing the efficiency of the optoelectronic surveillance system while minimizing the power consumed by active types of surveillance


DOI 10.15588/1607-3274-2024-4-15

equipment in the presence of boundary constraints on the efficiency and power consumed by the system.

The proposed indicator of the system's functioning efficiency allows us to assess the level of impossibility of uncontrolled crossing of the perimeter of the protected area by an intruder.

**The practical significance** of the obtained results lies in the programmatic implementation of the proposed mathematical apparatus for assessing the efficiency of OSSs functioning in different conditions under which observation is carried out.

**Prospects for further research** include solving the problem under study by choosing different types of surveillance equipment or their combinations, on different towers as well as testing the hypothesis that this may achieve better results in some cases.

## ACKNOWLEDGEMENTS

The work was carried out within the framework of research works № 6 "Methodology for assessing the effectiveness of construction of the state border protection outside border-crossing points in the area of responsibility of the state border protection unit", № 20 "Geoinformation support for state border protection", № 22 "Features of state border protection at the border guard detachment site in the conditions of active actions of enemy sabotage forces", № 35 "Methodology for processing calculations for the maintenance of border infrastructure in sections of the state border", approved by the Plan of Scientific and Technical Activities of the State Border Guard Service of Ukraine for 2024–2026 (Order of the Administration of the State Border Guard Service of Ukraine № 476-AH of 19.04.2024).

The work was positively evaluated by the Department of Communication and Information Systems and the Department of General Science and Engineering Disciplines of the Faculty of Operational Support of Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine.

## REFERENCES

1. Rachok R. V., Borovyk O. V., Borovyk L. V. Method of effective placement of various elements of a complex surveillance system, *Radioelectronics, informatics, control*, 2019, № 3(50), pp. 123–130.
2. Elharrouss O., Almaadeed N., Al-Maadeed S. A review of video surveillance systems, *Journal of Visual Communication and Image Representation*, 2021, Vol. 77, pp. 37–43.
3. Obod I., Svyd I., Maltsev O., Vorgul O., Maistrenko G. and Zavolodko G. Optimization of Data Transfer in Cooperative Surveillance Systems, *International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*. Kharkiv, 2018, pp. 539–542.
4. Zhang, Liangpei & Zhang, Lefei & Du, Bo Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art, *IEEE Geoscience and Remote Sensing Magazine*, 2016, pp. 22–40.
5. Li, Ying & Zhang, Haokui & Xue, Xizhe & Jiang, Yenan & Shen, Qiang Deep learning for remote sensing image classification: A survey, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2018, pp. 147–156.
6. Shakhatreh H., Sawalmeh A., Al-Fuqaha A., Dou Z., Almaitta E., Khalil I., Othman N., Khreishah A., Guizani, Mohsen Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges, *IEEE Access*, 2019, № 7, pp. 48–57.
7. Barannik V., Barannik D., Fustii V., Parkhomenko M. Evaluation of Effectiveness of Masking Methods of Aerial Photographs. 3rd International Conference on Advanced Information and Communications Technologies. Lviv, 2019, pp. 415–418.
8. Barannik V., Krasnoruckiy A., Larin V., Hahanova A. and Shulgin S. Model of syntactic representation of aerophoto images segments, *Modern Problems of Radio Engineering, Telecommunications and Computer Science, (TCSET'2018): XIVth Intern conf.* Lviv-Slavske, 2018, pp. 974–977.
9. Tarshin V. A., Tansyura O. B. , Kuravsky M. V. Ways to increase the probability of detection and recognition of objects on the images of multispectral optical-electronic reconnaissance systems, *Science and technology of the Armed Forces Forces of Ukrain*, 2021, № 1, pp. 141–146.
10. Tarshin V. A., Tansyura O. B., Kuravsky M. V. Ways to improve the informativeness of multi-spectral optical and electronic surveillance systems of unmanned aerial vehicles, *Science and technology of the Air Force of the Armed Forces of Ukraine*, 2023, № 1, pp. 56–62.
11. Borovyk O. V., Rachok R. V., Darmoroz M. M. Evaluation of the effectiveness of the operation of the optical-electronic surveillance system, *Radioelectronics, informatics, control*, 2017, № 2(41), pp. 93–99.
12. Rachok R. V., Borovyk O. V., Borovyk L. V. Structural optimization of the optical-electronic surveillance system, *Radioelectronics, informatics, control*, 2017, № 4(43), pp. 151–161.

Received 20.08.2024.
Accepted 09.11.2024.


DOI 10.15588/1607-3274-2024-4-15

УДК 528.29

# МЕТОДИКА ОПТИМІЗАЦІЇ ФУНКЦІОНУВАННЯ СИСТЕМИ ОПТИКО-ЕЛЕКТРОННОГО СПОСТЕРЕЖЕННЯ

**Боровик Д. О.** – Магістр з комп. наук, Хмельницький, Україна.
**Боровик О. В.** – д-р техн. наук, професор, заступник начальника відділу організації освітньої та наукової діяльності управління професійної підготовки Адміністрації Державної прикордонної служби України, Київ, Україна.
**Рачок Р. В.** – д-р техн. наук, професор, професор кафедри зв'язку та інформаційних систем, Національна академія Державної прикордонної служби України імені Богдана Хмельницького, Хмельницький, Україна.
**Басараба І. О.** – д-р філософії з філології, викладач кафедри іноземних мов, Національна академія Державної прикордонної служби України імені Богдана Хмельницького, Хмельницький, Україна.

## АНОТАЦІЯ

**Актуальність.** При охороні державного кордону активно застосовуються радіолокаційні, тепловізійні та відео засоби спостереження. Разом з відповідним комунікаційним обладнанням вони дозволяють утворювати системи оптико-електронного спостереження, які є основою для інтелектуалізації охорони кордону. На ефективність використання таких систем суттєво впливають особливості їх функціональної і структурної побудови. Раціональна структурна побудова навіть у складних фізико-географічних умовах дозволяє забезпечувати високий рівень ефективності спостереження. Однак значний вплив на підвищення ефективності системи має і функціональна складова. Функціонування елементів системи оптико-електронного спостереження в багатьох випадках відбувається в умовах обмежень щодо використання електроживлення. Такі обмеження визначають необхідність раціонального вибору режимів використання окремих типів засобів спостереження на окремих часових інтервалах з метою забезпечення ефективного спостереження з урахуванням періоду доби та погодних умов. Недосконаліть науково-методичного апарату оптимізації функціонування систем оптико-електронного спостереження визначає актуальність даного дослідження.

**Мета.** Метою роботи є розробка методики оптимізації функціонування системи оптико-електронного спостереження за рахунок раціонального вибору режимів функціонування різних типів засобів спостереження на окремих часових інтервалах з урахуванням періоду доби та погодних умов, в яких вони використовуються.

**Метод.** У роботі поставлена та досліджена двокритеріальна задача вибору режимів функціонування різних типів засобів спостереження системи оптико-електронного спостереження на окремих часових інтервалах з урахуванням періоду доби та погодних умов, в яких вони використовуються, що забезпечує максимізацію ефективності функціонування системи оптико-електронного спостереження при мінімізації потужності спожитої електроенергії активними типами засобів спостереження за наявності граничних обмежень щодо ефективності та потужності спожитої системою електроенергії.

Запропонований показник оцінки ефективності функціонування системи дозволяє оцінити рівень неможливості неконтрольованого перетину порушником периметру ділянки, що охороняється. Особливістю зазначеної методики є можливість забезпечення суттєвого зниження рівня енергоспоживання складовими системи за рахунок незначного зниження ефективності спостереження.

**Результати.** У роботі запропонований альтернативний підхід для оцінки ефективності функціонування системи оптико-електронного спостереження, ідея якого полягає в тому, що замість оцінювання ефективності спостереження по всій площі контрольованої області ділянки кордону, оцінюється ефективність контролю лише за периметром цієї області. Такий підхід суттєво зменшує обчислювальну складність задачі відшукання значення ефективності, що в подальшому спрощує вирішення задач структурної оптимізації систем спостереження. Запропоновано програмно-алгоритмічну реалізацію методики оптимізації функціонування системи оптико-електронного спостереження. З використанням розробленого програмного забезпечення проведено раціональний вибір режимів функціонування окремих типів засобів спостереження на окремих часових інтервалах з урахуванням періоду доби та погодних умов.

**Висновки.** Використання запропонованої методики дозволяє оптимізувати режими функціонування системи оптико-електронного спостереження з урахуванням обмежуючих факторів щодо ефективності та потужності енергоспоживання при застосуванні однакових типів засобів спостереження на всіх вежах системи. Можливим напрямом удосконалення методики є її адаптація до випадків застосування на різних вежах системи різних типів засобів спостереження.

**КЛЮЧОВІ СЛОВА:** система оптико-електронного спостереження, методика, оптимізація, ефективність, енергоспоживання, алгоритм.


DOI 10.15588/1607-3274-2024-4-15

**ЛІТЕРАТУРА**

1. Рачок Р. В. Методика ефективного розміщення різнотипних елементів складної системи спостереження / Р. В. Рачок, О. В. Боровик, Л. В. Боровик // Радіоелектроніка, інформатика, управління. – 2019. – № 3(50). – С. 123–130.
2. Elharrouss O. A review of video surveillance systems / O. Elharrouss, N. Almaadeed , S. Al-Maadeed // Journal of Visual Communication and Image Representation. – 2021. – Vol. 77. – P. 37–43.
3. Optimization of Data Transfer in Cooperative Surveillance Systems / [I. Obod , I. Svyd , O. Maltsev et al.] // International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). – Kharkiv, 2018. – P. 539–542.
4. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. / [Zhang, Liangpei & Zhang, Lefei & Du, Bo] // IEEE Geoscience and Remote Sensing Magazine. – 2016. – P. 22–40.
5. Deep learning for remote sensing image classification: A survey / [Li, Ying & Zhang, Haokui & Xue et al.] // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. – 2018. – P. 147–156.
6. Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges / [H. Shakhatreh , A. Sawalmeh , A. Al-Fuqaha et al.] // IEEE Access. – 2019. – № 7. – P. 48–57.
7. Evaluation of Effectiveness of Masking Methods of Aerial Photographs / [V. Barannik, D. Barannik , V. Fustii, M. Parkhomenko] // 3rd International Conference on Advanced Information and Communications Technologies. – Lviv, 2019. – P. 415–418.
8. Model of syntactic representation of aerophoto images segments / [V. Barannik , A. Krasnoruckiy , V. Larin et al.] // Modern Problems of Radio Engineering, Telecommunications and Computer Science, (TCSET’2018) : XIVth Intern conf. – Lviv-Slavske, 2018. – P. 974–977.
9. Таршин В. А. Шляхи підвищення імовірності виявлення та розпізнавання об’єктів на зображеннях різноспектральних оптико-електронних систем розвідки / В. А. Таршин, О. Б. Танцюра, М. В. Куравський // Наука і техніка Повітряних Сил Збройних Сил України. – 2021. – № 1. – С. 141–146.
10. Таршин В. А. Шляхи покращення інформативності різноспектральних оптико- електронних систем спостереження безпілотних літальних апаратів / В. А. Таршин, О. Б. Танцюра, М. В. Куравський // Наука і техніка Повітряних Сил Збройних Сил України. – 2023. – № 1. – С. 56–62.
11. Боровик О. В. Оцінка ефективності функціонування системи оптико-електронного спостереження / О. В. Боровик, Р. В. Рачок, М. М. Дармороз // Радіоелектроніка, інформатика, управління. – 2017. – № 2(41). – С. 93–99.
12. Рачок Р. В. Структурна оптимізація системи оптико-електронного спостереження / Р. В. Рачок, О. В. Боровик, Л. В. Боровик, // Радіоелектроніка, інформатика, управління. – 2017. – № 4(43). – С. 151–161.


DOI 10.15588/1607-3274-2024-4-15

UDC 681.326

# ADAPTIVE FILTERING AND MACHINE LEARNING METHODS IN NOISE SUPPRESSION SYSTEMS, IMPLEMENTED ON THE SoC

**Shkil A. S.** − PhD, Associate Professor, Associate Professor of Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

**Filippenko O. I.** – PhD, Associate Professor, Associate Professor of Infocommunication Engineering Department named by V.V. Popovsky, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

**Rakhlis D. Y.** − PhD, Associate Professor, Associate Professor of Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

**Filippenko I. V.** – PhD, Associate Professor, Associate Professor of Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

**Parkhomenko A. V. –** PhD, Associate Professor, Associate Professor of Software Tools Department, National University Zaporizhzhia Polytechnic, Zaporizhzhia, Ukraine.

**Korniienko V. R. –** PhD student of Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

## ABSTRACT

**Context.** Modern video conferencing systems work in different noise environments, so preservation of speech clarity and provision of quick adaptation to changes in this environment are relevant tasks. During the development of embedded systems, finding a balance between resource consumption, performance, and signal quality obtained after noise suppression is necessary. Systems on a chip allow us to use the power of both processor cores available on the hardware platform and FPGAs to perform complex calculations, which contributes to increasing the speed or reducing the load on the central SoC cores.

**Objective**. To conduct a comparative analysis of the noise suppression quality in audio signals by an adaptive filtering algorithm and a filtering algorithm using machine learning based on the RNNoise neural network in noise suppression devices on the technological platform SoC.

**Method**. Evaluation using objective metrics and spectrogram analysis using the Librosa library in Python. Neural network training and model design are performed on the basis of Python and Torch tools. The Vitis IDE package was used for the neural network implementation on the platform SoC.

**Results**. The analysis of two noise suppression methods using the adaptive Wiener filter and the RNNoise neural network was performed. In the considered scenarios, it was determined that the neural network shows better noise suppression results according to the analysis of spectrograms and objective metrics.

**Conclusions.** A comparative analysis of the effectiveness of noise suppression algorithms based on adaptive filters and a neural network was performed for scenarios with different noise environments. The results of objective SIGMOS metrics were obtained to evaluate the quality of the received audio signal. In addition, the possibility of running the RNNoise neural network on the technological platform SoC ZYNQ 7000 was verified.

**KEYWORDS:** embedded systems, system-on-a-chip, FPGA, adaptive filtering, digital signal processing algorithms, noise suppression algorithms, audio signals, machine learning, neural networks.

## ABBREVIATIONS

ARM is an advanced RISC machine;
CNN is a convolutional neural network;
DNN is a deep neural network;
FPGA is a field programmable gate array;
FPS is frames per second;
GRU is a gated recurrent unit;
GTCRN is group temporal convolutional recurrent network;
GZIP is a GNU's not unix ZIP;
HLS is a high level synthesis;
ISP is an ITU;
ITU is an international telecommunication union;
LPF is a low pass filter;
MOS is a mean opinion score;
PESQ is a perceptual evaluation of speech quality;
POLQA is a perceptual objective listening quality analysis;
PSTN is a public switched telephone network;
RNN is a recurrent neural network;
SNR is a signal to noise ratio;
SoC is a system on a chip;
VAD is a voice activity detector;
VISQOL is a virtual speech quality objective listener;
VoIP is a Voice over Internet Protocol.

## NOMENCLATURE

$a$,$b$ are coefficients of the biquadratic filter;
$D_{sig}$ is a set of signals to be investigated;
$h(n)$ is a set of the corresponding coefficients for the given filter;
$i$ is an index of the input signal;
$j$ is an index of the noise sample;
$k$ is an index of the frequency domain bin;
$k(n)$ is a noisy audio signal with $n$ samples in it;
$L$ is a number of signals to be tested;
$M$ is a length of the window function;
$N$ is a length of the input signal buffer;
$n$ is a index of count/sample;
$S$ is a maximum number of counts in the signal;


DOI 10.15588/1607-3274-2024-4-16

$w(k)$ is a window function;

$X[k]$ is a frequency domain representation of the signal;

$x(n)$ is an input signal with speech and noise environment;

$x_c(n)$ is an useful signal with $n$ samples in it;

$y(n)$ is an output signal.

## INTRODUCTION

Currently, corporate video conferencing systems are actively evolving, where the determining factors are ensuring the reliability of information transmission, improving the quality of the transmitted far-end and near-end audio and video parts, and ensuring the reliability of the communication device.

In the field of audio signal processing, one of the improvement options of the quality and analysis is adaptive filtering usage, which solves a wide range of tasks. In particular, typical tasks are noise suppression with the selection of human speech, dereverberation and echo canceling, and signal separation. The different nature of the suppressed noises complicates the design of adaptive filters due to the need to constantly update the coefficients and adjust to a specific type of noise.

In particular, it is worth noting that different noise environments can have combined types of noise (stationary and non-stationary), which makes the task of adaptive filter design even more challenging.

Examples of such noise environments and their combinations are office or street noise, industrial rooms with ventilation. These types of noise are examples of stationary noise and its combination with momentary bursts in the spectrum, which can be from a keyboard (in the case of an office) or different sources that have a similar spectrum to human speech. Due to the different nature of noise, the use of classical adaptive filtering algorithms does not give a positive result for the user.

Machine learning algorithms based on neural networks, which can determine the type of noise and perform both Voice Activity determination (the presence of a voice in an audio fragment) and noise suppression in the resulting audio stream, have gained popularity in the tasks of noise classification and separation. The use of neural networks is a complex computational task that cannot be solved on stationary hardware due to the need for both a specialized software environment and compliance with the time requirements set by the real time audio streaming industry.

The main feature of using neural networks for adaptive noise suppression is the ability to perform training on a specific sample of noises, which improves performance in a specific environment. Typical noise suppression libraries for audio processing are RNNoise and Speexdsp.

Deep learning is used in noise suppression tasks for several key reasons, mainly because it allows significantly improve results compared to traditional signal processing methods. Traditional noise suppression methods are often based on linear models or statistical assumptions about noise characteristics. Deep neural networks are able to model complex non-linear dependencies between pure signal and noise, which allows to more efficiently separate the useful signal even in difficult conditions. Deep learning models can adapt to different types of noise and signals.

They can be trained on large data sets containing examples of different recording conditions and noise types, which makes them universal in use. Unlike traditional methods, which often require the manual design of features for each specific application, deep neural networks can automatically extract relevant features from raw data. It dramatically reduces the need for feature engineering and allows the creation of more general models.

Thus, **the object of the study** is adaptive filtering procedures in digital signal processing systems.

**The subject of the study** are models, methods and procedures for designing adaptive digital filters on the technological platform SoC using machine learning methods.

**The objective of the study** is to perform a comparative analysis of the noise suppression quality in audio signals using an adaptive filtering algorithm and a filtering algorithm using machine learning based on the RNNoise neural network in noise suppression devices on the technological platform SoC.

## 1 PROBLEM STATEMENT

Due to the above features and requirements, one of the options to solve the problem of adaptive noise suppression is the implementation of the specified algorithms on a specialized hardware platform from the SoC family. This provides an opportunity both to implement the business-logic of the firmware, executed on the device, and to design and use specialized hardware accelerators to ensure the optimal distribution of the data processing path in the SoC, taking into account limited resources both for processing time, and for maximum system power consumption as a whole.

Let's use a set of audio signals containing different types of noise and human speech. Each input signal from the set includes a noise environment in a given ratio to human speech. Thus, the signal-to-noise ratio and the present noise environment type characterize each element of the set.

Let's denote a set of audio signals as $D_{sig}$ where each element is characterized by a noise environment and its SNR.

The data set can be expressed as $D_{sig} = \{x_i(n)\}$, where $i = 1,2,\ldots,L$ – the number of signals, $n = 0,1,2,\ldots,S$ – the number of counts of the corresponding signal. Each $i$-th signal is a combination of the useful signal $x_{ci}(n)$ and the $j$-th noise environment $k_j(n)$ in a given signal-to-noise ratio.

Thus, the input sample of signals makes a set of combinations $x_i(n) = \langle x_{ci}(n), k_j(n) \rangle$, where $i$


DOI 10.15588/1607-3274-2024-4-16

corresponds to the index of the speech signal, $j$ corresponds to the index of the noise environment signal, $x_i(n)$ corresponds to their combination with the given SNR. In every noise suppression scenario, it is necessary to suppress $k_j(n)$ as much as possible and obtain a clean output signal $y_i(n)$. According to the general filtering equation $y_i(n) = h_i(n) \cdot x_i(n)$, the final result of using both a neural network and an adaptive filter is to find the coefficients $h_i(n)$ for each of the combinations of noise environment and speech during the processing of the input audio signal.

The main task of adaptive filtering for noise suppression is the quick response of the filter to changes in the noise environment, preservation of the dynamic range of the input signal, and suppression of signal fragments where there is no human speech, i.e., quick selection of coefficients. Several filtering algorithms are designed for noise suppression, among which the adaptive Wiener filter, LMS filter, and RNNoise neural network are known.

At the same time, the output signal after filtering is characterized by a set of objective metrics of audio signal quality, such as discontinuity, noiseiness, reverberation, coloration, and overall quality assessment, which can be used for the quality characteristics of the noise suppression algorithm.

When using the RNNoise neural network, a 10-millisecond-long input signal buffer at a sampling frequency of 48 KHz must be provided to its input; that is, the input buffer should contain 480 counts of the input signal.

The main criterion for comparing different noise suppression algorithms is the analysis of the obtained objective metrics of the output signal and visual observation of audio signal spectrograms before and after noise suppression.

Thus, the tasks of the study are the selection of an adaptive filtering algorithm and a machine learning method based on a neural network for noise suppression, the determination of metrics that can be used to evaluate the result of noise suppression, and the performance of a comparative analysis of the selected methods.

## 2 REVIEW OF THE LITERATURE

Paper [1] considers the possibility of using primitives of modern C++ standards for high-level synthesis. Features of the latest language standards and their features during synthesis are given and considered. A comparative analysis of Catapult HLS with Vitis HLS in terms of support for modern C++ primitives during synthesis was conducted. The hardware-accelerated matrix determinant calculator was implemented using C++03 and C++17.

Paper [2] considered the features of using high-level synthesis during the development of image processing algorithms. The main points related to synthesized and non-synthesized constructions of different variants of C-code to get maximum performance are given. As an example, the implementation of the Sobel filter algorithm with the analysis of performance changes from 10 to 388 FPS was used. Analysis of changes of IP-core main code to obtain a necessary performance was done.

Paper [3] analyzes the process of developing a cluster to implement GZIP compression using the ZedBoard debugging board. A comparative analysis of the energy efficiency of the obtained implementations and the productivity of the obtained cluster was conducted based on the Wordcount and Terasort benchmarks. According to the conclusions, a two-fold improvement in the energy consumption reduction of the cluster was obtained.

In modern models the speech enhancement based on deep learning has made significant strides compared to traditional methods, but it often requires a large number of parameters and significant computing power, which makes them difficult to use on resource-constrained devices in real conditions. In paper [4], a GTCRN is considered, which uses group strategies to effectively simplify the competitive model of Dual-Path Convolution Recurrent Network. In addition, it applies subband feature extraction modules and temporal recurrent models to improve performance. The obtained model requires low computational resources, having only 23.7 thousand parameters and 39.6 million accumulation operations per second. Experimental results show that the proposed model not only outperforms RNNoise, a typical lightweight model with similar computational load, but also exhibits competitive performance compared to current baseline models that require significantly more computing resources.

Paper [5] proposes an accelerator for the Markov decision process, implemented in the AI-Toolbox public library using high-level synthesis tools using the “tiger-antelope” problem as an example. The proposed approach shows an acceleration of more than 7 times compared to the original version of the algorithm.

Hardware accelerators for deep learning algorithms is a relevant topic in the scientific and engineering environment. The use of FPGAs as embedded co-processors to accelerate algorithms has gained wide recognition.

Decision-making games are critical tasks where their real-time accuracy and efficiency directly affect the result. Traditional FPGA development using hardware description languages is associated with long development cycles, high complexity, slow iterations, and problems in fast responding to model algorithm updates. In comparison, HLS-based design of FPGA provides an appropriate technology path to eliminate these weaknesses.

Paper [6] considers the FPGA usage together with a high-level synthesis to solve problems in the field of decision-making algorithms design. A comparative analysis of the algorithm implementation on the FPGA with the software solution based on the x86 architecture is carried out. A dedicated deep learning algorithm related to decision making games was implemented, optimized and deployed on FPGA using HLS. The FPGA is


DOI 10.15588/1607-3274-2024-4-16

connected to the host PC via PCIe as a co-processor. Comparative testing with running the algorithm on an eight-core ftd2000 CPU and an Intel i9-9900k processor demonstrates that using the FPGA as a co-processor significantly reduces the execution time of the algorithm, resulting in a noticeable speedup effect.

Paper [7] presents an experimental study of the implementation of a face mask detection system based on the use of HLS and the concept of hardware and software co-design. The target platform is a Xilinx PYNQ-Z2 FPGA board that connects to the host computer and acts as a hardware accelerator for the face masks detection task. To simplify the hardware implementation complexity, the face mask detection algorithm uses the ISP approach instead of complex CNN models. The algorithm consists of color space transformation, skin color detection, morphological operations, connected component labeling and horizontal edge detection. Implementation results show that the FPS for detection can reach 18.

Deep neural networks are widely used to solve a variety of tasks: from speech recognition to image classification. Since DNNs require a lot of computing power, their hardware implementation on FPGA or ASIC has attracted significant attention. In turn HLS is widely used because it significantly increases performance and flexibility and requires minimal hardware knowledge. Paper [8] proposes DeepFlexiHLS, a two-stage design space exploration flow for searching a set of directives to achieve minimum latency. The results form a Pareto space from which the developer can choose whether his FPGA resources are limited or should not be fully utilized by the DNN module.

The growing trend of using deep learning techniques for noise suppression has led to the creation of hybrid noise suppression systems that combine classical signal processing with deep learning. For example, paper [9] focuses on expanding the RNNoise noise suppression system by including additional features during the training phase. The paper presents a detailed description of the configuration process of the modified system and the comparative results obtained from the performance analysis using the existing version of RNNoise as a benchmark.

Deep learning-based speech enhancement can provide near-best performance when processing non-stationary noise. Noise suppression methods, that combine classical signal processing with a RNN, can be implemented in real time due to their low complexity. However, in these methods long-term speech information is missed during features selection, which degrades the performance of noise suppression. This paper extends the RNN-based denoising method known as RNNoise by adding a long-term spectral difference feature. The amount of noise suppression is also limited to improve speech quality for a better compromise between noise removal level and speech distortion. The method proposed in [10] outperforms the RNNoise algorithm by 0.12 MOS points on average according to the results of the subjective audio test.

Background noise is a major source of quality degradation in VoIP and PSTN calls. Recent studies have shown the effectiveness of deep learning for noise suppression, but the datasets used were relatively small compared to other domains (e.g., ImageNet) and the corresponding evaluations were more concentrated. To better support deep learning speech enhancement research, paper [11] presents a noisy speech dataset that can scale to arbitrary sizes based on the number of speakers, noise types, and SNR levels. Increasing the dataset size shows the improvement of the noise suppression performance as expected. To demonstrate the data set and evaluation structure, subjective MOS with objective quality measures such as SNR, PESQ, POLQA, and VISQOL, was applied, and reasons of MOS relevance were demonstrated.

In teleconferencing scenarios, the speech is usually degraded by background noise, which reduces speech intelligibility and quality. Therefore, it is extremely important to improve speech in noisy environments. Paper [12] investigates a real-time speech enhancement method for a far-end signal based on an improved RNN with a unit of controlled GRU. The ideal amplitude masking values of the reverberated target speech are used as the target values for RNN training. Feature normalization is applied and subband normalization technology were proposed to reduce feature differences, which contributes to better RNN learning of long-term patterns. In addition, to suppress the residual interharmonic pseudo-steady noise due to subbanding, the work integrates RNN with the optimal modified log-spectral amplitude algorithm. Experimental results show that the proposed method improves speech quality and reduces distortion with low computational complexity for real-time operation.

Paper [13] considers the implementation of the Librosa software package for the analysis of audio data and operations on it. The Librosa library is widely used in research and development in digital signal processing and machine learning for audio data processing. The main capabilities and functions of Librosa are considered, such as loading and saving audio files, spectrogram calculation, mel-spectrograms and mel-frequency cepstral coefficients, as well as methods for tonal analysis, segmentation and extraction of sound features. Special attention is paid to functions for pre-processing of audio signals, including normalization, noise reduction and changes in playback speed.

Paper [14] considers the implementation of the Pyroomacoustics package, designed for rapid development and testing of audio signal processing algorithms using microphone arrays. The package contains three main components: an object-oriented interface on Python that allows quickly create various simulation scenarios involving multiple sound sources and microphones in 2D and 3D rooms; a fast C-implementation of a source image model for general


DOI 10.15588/1607-3274-2024-4-16

multifaceted rooms that efficiently generates room impulse characteristics and models sound propagation between sources and receivers; reference implementations of popular algorithms for beamforming, direction determination and adaptive filtering. Together, these components form a package that significantly reduces the time to implement new algorithms, reducing overhead at the performance evaluation stage.

Paper [15] considers the problem of single-channel speech enhancement in stationary conditions and proposes the use of a Wiener filter with a recursive noise estimation algorithm. The Wiener filter is a linear estimator and minimizes the root mean square error between the original and enhanced speech. The algorithm is implemented in the frequency domain and it depends on the transfer function of the filter, which varies from sample to sample based on the statistics of the speech signal, in particular the local mean and local dispersion. A recursive noise estimation approach is used for noise estimation. In this approach, noise estimation is performed based on past and present spectral power values using a smoothing parameter. The value of the smoothing parameter is chosen in the range from 0 to 1. To evaluate the effectiveness of the proposed speech enhancement algorithm, objective evaluations and informal listening of sentences from the NOIZEUS corpus, spoken by male and female voices, distorted by white and pink noise at different levels of the signal/noise, are conducted. SNR, segmental SNR and perceptual assessment of speech quality are used for objective measurements. The measurement results prove that speech, enhanced by the proposed algorithm, is more pleasing to the human ear under both noise conditions, compared to traditional speech enhancement methods.

## 3 MATERIALS AND METHODS

In the tasks of digital signal processing, the classic solution for noise suppression is the use of adaptive filters, which mostly work on the principle of minimizing the root mean square error between the signal with a noise component and the restored signal. Variations of Wiener adaptive filter [15] can include the analysis of harmonics in the signal, the use of VAD to determine the presence of speech. The main disadvantages of using adaptive filters are the need to determine the exact noise model and its stationarity. In many real-world situations, these characteristics can be unknown or difficult to estimate. In real conditions, signals and noise are often non-stationary, which can significantly reduce the effectiveness of the adaptive filter.

This study uses the RNNoise neural network, which is a RNN designed to work in environments with limited computing power. Neural network training and model design is performed on the basis of Python and Torch tools on a PC with x86 architecture, after which the interference (launch) of the model is performed in the environment, developed on C. The interference part of the trained model is performed on the ARM part of the SoC ZYNQ 7000 using the appropriate set of tools for bare-metal launch. The architecture of the RNNoise neural network is shown in Fig. 1.

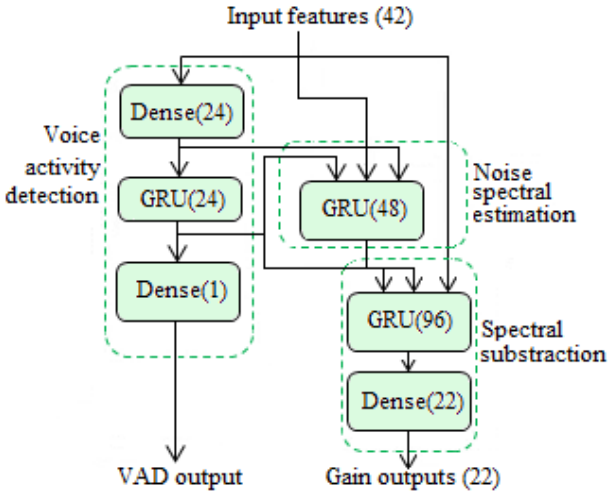



Figure 1 – General architecture of RNNoise neural network

Each block represents a layer of neurons, the number of which is indicated in parentheses. Dense layers: dense (24) and dense(1), are fully connected layers which are not repeated. One of the network outputs is a set of gains that are applied at different frequencies (gain outputs (22)). Another result is the probability of the voice activity that is not used for noise suppression, but is a useful by-product of the network, which can be further used to implement automatic signal level control after noise suppression. According to this strategy, the level equalization is performed taking into account the VAD data, which allows for additional noise suppression between speech fragments.

Note that, in general, filtering due to the use of an adaptive filter can be represented by the equation (1):

$$y(n) = h(n) \cdot x(n). \quad (1)$$

At the same time, in the case of using the RNNoise neural network, stages related to the preparation of data for processing by the neural network and data conversion from the frequency range to the time range are added to the signal processing path. RNNoise performs processing on audio frames equivalent to 10 milliseconds of audio stream with a sampling frequency of 48KHz. Signal processing includes a biquad filter, convolution with a window function and fast Fourier transform computation.

In general, the process of preprocessing of the input signal can be expressed as follows. Let's consider an input signal buffer that contains 10 milliseconds of input data stream. In general, the input signal buffer can be represented as a one-dimensional vector with a given number of elements: $X=[x(0), x(1), x(2),\ldots, x(N–1)]$.

If the input signal is represented as a vector $x(t)$ and the biquadratic filter has coefficients $b = [b_0, b_1, b_2]$ and $a = [1, –a_1, –a_2]$, where $a_0$ is usually equal to 1, then


DOI 10.15588/1607-3274-2024-4-16

signal filtering using a biquadratic filter can be defined as (2):

$$x_b[n] = b_0 x[n] + b_1 x[n-1] + b_2 x[n-2] - a_1 y[n-1] - a_2 y[n-2], \quad (2)$$

where $y[n]$ – the output value of the filter, $b(0\ldots2)$ and $a(1\ldots2)$ – filter coefficients, $x[n]$ – the value of the input signal at the corresponding time.

In turn, the next stage with the convolution of the signal with a window function can be defined as (3):

$$y[n] = \sum_{k=0}^{M-1} h(n-k) \cdot w(k). \quad (3)$$

And accordingly, the fast Fourier transform using the Kiss_fft library is generally defined as (4):

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-\frac{2\pi kn}{N}}. \quad (4)$$

After that, signal analysis in the frequency domain, band gain calculation for opus frequency bands and inverse Fourier transformation are performed.

Several examples of signal and noise combinations were chosen for the study, namely road noise, street noise and crowd noise. The SNR was chosen to be 5.10 dB and 15dB, respectively.

Processing with the help of the Wiener filter was performed using the Pyroomacoustics package and its module Denoise [14].

## 4 EXPERIMENTS

For effective study of noise suppression with the help of the selected algorithms, the test signals of crowd noise, road noise and street noise were selected in ratios with the useful signal SNR5/SNR10/SNR15.

Spectrograms of some of the received resulting signals and a comparative table of all conducted experiments are given in the paper. As an example for further consideration, we will use an input signal composed of fragments of male and female speech with intervals between replicas about 0.8 seconds. The spectrogram of the input signal without noise is shown in Fig. 2.

An experimental study of the obtained processing speed was performed on the hardware platform ZYNQ Zedoard [16]. The analysis of spectrograms and the obtained results of noise suppression in the study were performed with the help of developed utilities based on libraries Librosa, Libsndfile, Pyroomacoustics, Pyloudnorm, Numpy and Matplotlib for the development and analysis of DSP algorithms on an ARM-based PC.

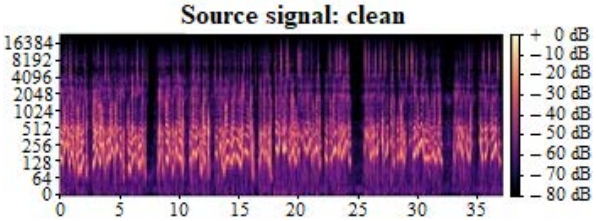



Figure 2 – Spectrogram of the input signal without noisy environment

To obtain objective signal metrics before and after processing, the SIGMOS package [17, 18] was used to determine discontinuity, noisenes, reverberation, coloration, and overall quality assessment. Objective metrics allow us to determine quickly the quality of the received signal processing results and compare algorithm implementations with each other.

During the research, objective metrics were obtained for the input signal, for the signal with noise suppression based on the adaptive Wiener filter, and for the signal with noise suppression based on the RNNoise neural network. SIGMOS metrics are calculated using a neural network that was developed to evaluate echo and noise suppression algorithms in telecommunications environments. The quality of the received signal is measured in the SIG according to ITU-T P.804 (subjective diagnostic test method for conversational speech quality analysis). SIGMOS evaluates sound quality parameters according to P.804. This model was trained using subjectively annotated data from P.804 to simulate human perception of sound quality.

## 5 RESULTS

For each example of an input signal and noise combination, objective metrics were obtained and the resulting spectrogram of the signal was analyzed before and after noise suppression. Taking into account the characteristics of the subject field, it is appropriate to consider the ratio of SNR5 and SNR15 between the useful signal and the noisy environment.

In the following experiments, noise of a certain type (mixed/combined) will be added to this signal and it will be suppressed by different types of filters.

The following diagrams show the results of the obtained objective metrics for the input and cleaned signal, obtained using Librosa. The order of the spectrograms for each test scenario is as follows: the input signal, the signal after processing by the Wiener filter, and the signal processed by the neural network.

Fig. 3 shows the input speech signal with added road noise with the ratio SNR5 (Fig. 3a). Second spectrogram shows the result of noise suppression using the Wiener filter (Fig. 3b), and the third shows noise suppression using a neural network (Fig. 3c).


DOI 10.15588/1607-3274-2024-4-16

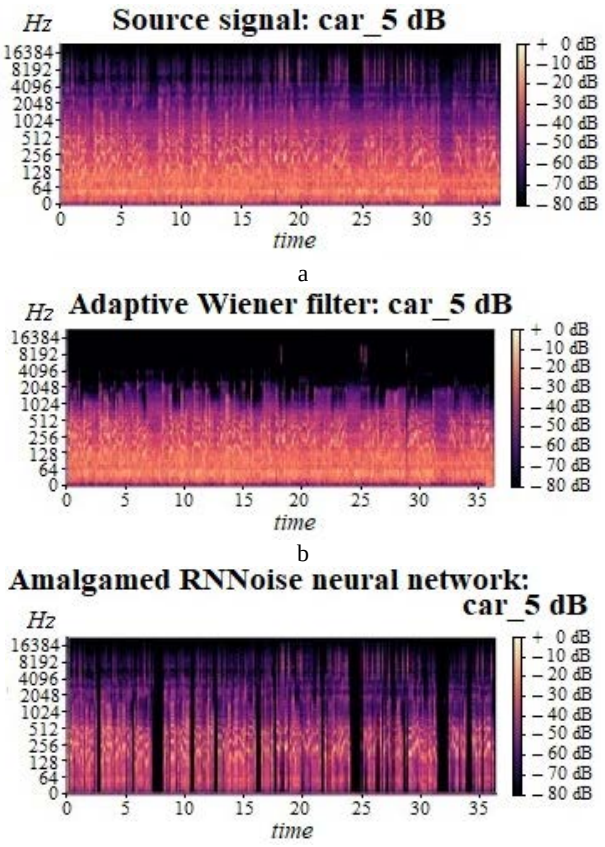



Figure 3 – Spectrograms of road noise suppression results with SNR5: a – input signal; b – noise suppression using the Wiener filter; c – noise suppression using neural network

The analysis of the second spectrogram (Fig. 3b) shows that the Wiener filter in this case worked as a LPF and slightly suppressed the fragments where there is no speech. Spectrum leaks between 15 and 30 seconds in the form of short-term high-frequency bursts are also visible.

During the analysis of the results of the objective metrics calculation, it was found out that absence of noise (namely Noiseness), is the best in the case of using a neural network. It is advisable to display the results of the SIGMOS calculation in the form of a histogram.

The data will be grouped into groups of three elements. Each group contains the results of one of the calculated MOS parameters for the input signal, the signal after processing using the Wiener filter and for the signal after processing with the RNNoise neural network. The result of objective metrics evaluation SIGMOS for the selected scenario is shown in Fig. 4.

The next stage was the analysis of noise suppression with a signal-to-noise ratio of 15dB. In this case, the neural network also performed better, which is evident from the given comparative spectrograms and objective metrics, obtained for input and output audio streams. The spectrograms of the input signal, the signal after cleaning using a Wiener filter and a neural network are represented in Fig. 5.

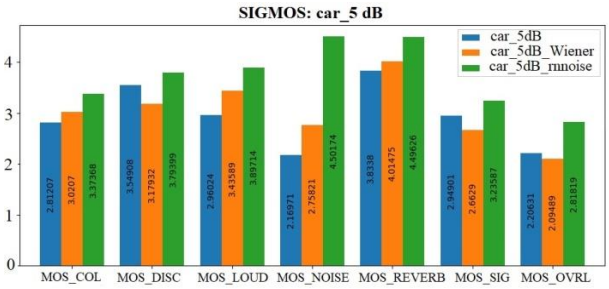



Figure 4 – SIGMOS objective metrics obtained for the corresponding human speech scenario with added road noise with SNR5

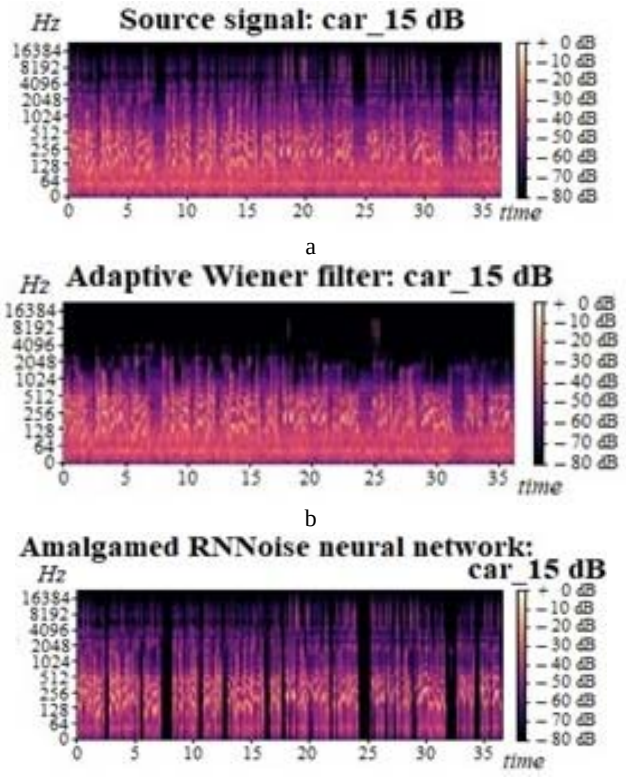



Figure 5 – Spectrograms of the the input signal for the scenario of human speech with added road noise with SNR15: a – input signal; b – noise suppression using the Wiener filter; c – noise suppression using neural network

It is appropriate to define that fragments are clearly visible, where there is no speech and only the noise component is present on the spectrogram after processing by the neural network. In the objective metrics for this recording, a trend with the best result of noise suppression using RNNoise, similar to the previous one, is observed. However, it should be noted that a small variation is present in the results for Loudness and Discontinuity between the input and output signals. The obtained SIGMOS objective metrics for the corresponding scenario with added road noise is shown in Fig. 6.

Next, a noise suppression analysis was performed for the scenario of added street noise to the speech with SNR5 and SNR15.


DOI 10.15588/1607-3274-2024-4-16

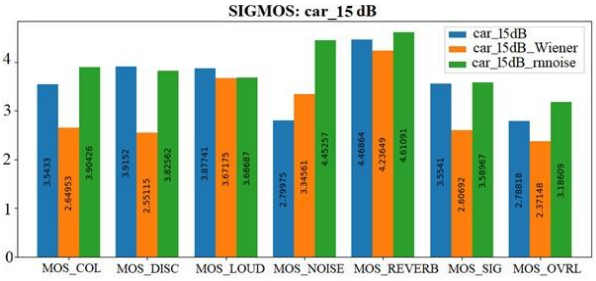



Figure 6 – SIGMOS objective metrics obtained for the corresponding human speech scenario with added road noise with SNR15

During the result analysis of obtained spectrograms, a similar trend of the Wiener filter in suppressing the high-frequency component is visible, but without detecting fragments where there is no speech. According to the obtained objective metrics, the neural network shows the best results in this scenario, compared to the Wiener filter. Fig. 7 shows the received spectrograms for the input signal, for the result of Wiener filter and RNNoise neural network processing. The results of objective metrics calculation for the scenario with added street noise with SNR5 is shown in Fig. 8.

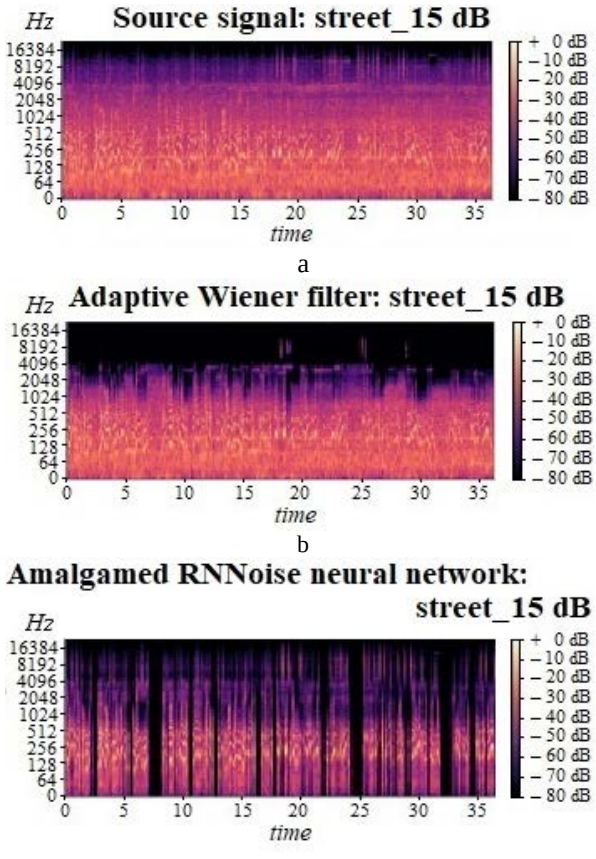



Figure 7 – Spectrograms of the obtained results for the scenario with added street noise with SNR5: a – input signal; b – noise suppression using the Wiener filter; c – noise suppression using neural network

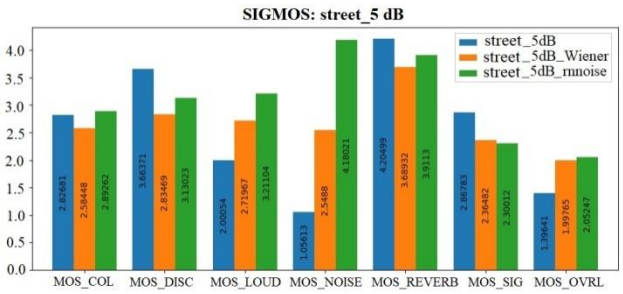



Figure 8 – SIGMOS objective metrics obtained for a human speech scenario with added street noise at SNR5

In the result of the analysis, tables, respectively, for the adaptive Wiener filter and the RNNoise neural network using Pandas tools were formed.

The obtained statistics of the objective metrics SIGMOS according to the experiment are shown in Table 1 and Table 2.

Metrics indicate the following characteristics of the received signal:

– MOS_COL (or MOS_COLORATION) is an indicator of audio coloration distortion. In the context of audio, coloration refers to a change in the sound timbre that can occur due to certain processing or recording/playback conditions;

– MOS_DISC (or MOS_DISCONTINUITY) is an indicator that characterizes the presence and degree of audio signal continuity;

– MOS_ LOUD (or MOS_LOUDNESS) is an indicator that characterizes the volume level of an audio signal. It evaluates as much as an audio volume matches the expected or desired level;

– MOS_NOISE is an indicator that evaluates the level of noise in an audio signal. It determines the degree of presence of background noise in the audio stream;

– MOS_REVERB (or MOS_REVERBERATION) is an indicator that characterizes the level of reverberation in an audio signal;

– MOS_SIG is an indicator that characterizes the quality of the audio signal itself without taking into account noise, reverberation or other distortions;

– MOS_OVRL (or MOS_OVERALL) is an overall indicator that reflects the overall quality of the audio, taking into account all the main aspects such as sound clarity, presence of noise, reverberation, signal continuity and other factors.

This is an integral indicator that combines all the other individual scores (MOS_COL, MOS_DISC, MOS_LOUD, MOS_NOISE, MOS_REVERB, MOS_SIG) to obtain a single value that reflects the overall quality of the audio signal.

SIGMOS/MOS has an absolute scale from 0 to 5, where a higher value indicates a better result for the selected metric, for example, for MOS_NOISENESS a higher value is a sign of the absence/suppression of noise.


DOI 10.15588/1607-3274-2024-4-16

Table 1 – Statistics of results for the adaptive Wiener filter

| | MOS_COL | MOS_DISC | MOS_LOUD | MOS_NOISE | MOS_REVERB | MOS_SIG | MOS_OVRL |
|---|---|---|---|---|---|---|---|
| count | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| mean | 2.648 | 2.974 | 3.178 | 2.939 | 3.764 | 2.687 | 2.214 |
| std | 0.246 | 0.407 | 0.332 | 0.603 | 0.326 | 0.185 | 0.195 |
| min | 2.178 | 2.436 | 2.720 | 2.178 | 3.224 | 2.365 | 1.998 |
| 25% | 2.56 | 2.648 | 2.89 | 2.579 | 3.586 | 2.583 | 2.079 |
| 50% | 2.606 | 2.891 | 3.182 | 2.842 | 3.668 | 2.687 | 2.155 |
| 75% | 2.836 | 3.288 | 3.420 | 3.096 | 3.988 | 2.751 | 2.309 |
| max | 3.021 | 3.674 | 3.672 | 4.359 | 4.236 | 3.046 | 2.651 |

Table 2 – Statistics of results for the RNNoise neural network

| | MOS_COL | MOS_DISC | MOS_LOUD | MOS_NOISE | MOS_REVERB | MOS_SIG | MOS_OVRL |
|---|---|---|---|---|---|---|---|
| count | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| mean | 3.514 | 3.715 | 3.691 | 4.347 | 4.377 | 3.208 | 2.839 |
| std | 0.432 | 0.388 | 0.282 | 0.209 | 0.313 | 0.557 | 0.543 |
| min | 2.893 | 3.130 | 3.211 | 4.026 | 3.911 | 2.300 | 2.052 |
| 25% | 3.219 | 3.552 | 3.526 | 4.199 | 4.076 | 2.920 | 2.561 |
| 50% | 3.454 | 3.738 | 3.735 | 4.344 | 4.485 | 3.265 | 2.845 |
| 75% | 3.766 | 3.883 | 3.905 | 4.453 | 4.589 | 3.383 | 2.965 |
| max | 4.350 | 4.478 | 4.077 | 4.757 | 4.800 | 4.348 | 4.049 |

The 25% and 75% values correspond to the first (Q1) and third (Q3) quartiles, respectively. These quartiles are part of descriptive statistics that show the distribution and central tendency of the data.

25% (first quartile, Q1) represents the value below which 25% of the data lies. This is the median of the lower half of the data set. This metric is useful for understanding the lower bound of the central 50% of the data.

75% (third quartile, Q3) represents the value below which 75% of the data lies. This is the median of the upper half of the data set, which is useful for understanding the upper bound of the central 50% of the data.

According to the Tables 1, 2, it is noticeable that the average value for all metrics of the neural network is higher, as well as both minimum and maximum. It is also noticeable that the noise suppression metric MOS_NOISE is 1.5 times higher than the results after filtering with an adaptive Wiener filter, which proves the effectiveness of using a neural network to suppress noise in different noisy environments, as well as MOS_COL, which indicates the coloration of the signal and the absence of tonal distortions.

Also, as part of the study, the RNNoise neural network was launched on the ARM part of the hardware platform ZYNQ. In particular, the necessary software finalizations of the source code were determined for the compilation of the interference part for the bare-metal environment. With the use of cross-compilation tools and refinement of the source code, the model execution time on the hardware platform was obtained, which allows us to perform the next stage of its performance analysis in the case of the development of a separate IP block using the Vitis-HLS toolkit.

## 6 DISCUSSION

Within the framework of this study, two approaches to filter the noise in an audio signal based on SoC were considered: adaptive filters and neural networks. Models with different bursts at random moments of time (non-stationary noise models) were used as a noise environment. Implemented adaptive filter and neural network for filtering signals with non-stationary noise and human speech in a given ratio were verified. The simulation results showed that both methods have significant potential for solving filtering problems, but show different results in terms of evaluation criteria.

Adaptive filters (the Wiener filter was considered) are relatively easy to implement (C, VHDL/Verilog), are effective for uniform noise, echo suppression due to the ability to achieve stable signal filtering without significant degradation in a relatively small number of counts. But, if the signal contains a non-stationary component (which is usually the real conditions of use), then the efficiency of the adaptive filter is significantly reduced.

A neural network can be implemented programmatically using both built-in accelerators, which are often found in SoCs, and with partial transfer of complex computing components to FPGA. Simulations have shown that neural networks can handle complex nonlinear dependencies much better than adaptive filters. For example, in the tasks of road and street noise filtering, where the noises had a complex structure, the RNNoise network showed a significantly better filtering result according to the objective metrics and the obtained spectrograms. Neural networks are effective in processing real-time changing signals and can also extract important features of the signal, which can reduce the need for manual tuning and data pre-processing unlike an adaptive filter. However, their implementation requires more computing resources and training time.


DOI 10.15588/1607-3274-2024-4-16

The results of the study show that the choice between adaptive filters and neural networks depends on the nature of the task. In systems, where stationary noise types predominate, adaptive filters are an appropriate choice. Their simplicity, transparency and adaptability make them suitable for processing audio signals with a relatively simple noise structure. Neural networks are appropriate for tasks where there are complex, non-linear noises and a high quality of noise suppression with maximum preservation of the useful signal is necessary. Also, their implementation in the SoC is possible due to the potential distribution of calculations between system components and due to the presence of hardware calculations accelerators, specific for neural networks, in some SoCs.

Also, it is necessary to consider the possibility of using during real-time processing of the audio stream objective metrics calculation and parallel processing in the situation, where combined approaches are possible, that is using adaptive filters and neural networks together to achieve maximum efficiency. For example, hybrid systems can use adaptive filters for primary signal processing, and neural networks – for more accurate filtering of complex noises in the second stage. Further research can be aimed at neural networks optimization for real-time operation and reduction of their requirements for computing resources and data volume.

Thus, the final choice between the specified filtering methods should be based on a detailed analysis of the specific task requirements, available computing resources, and power consumption limitations.

## CONCLUSIONS

During the study two methods of noise suppression using the adaptive Wiener filter and the RNNoise neural network, considered in different noisy environments, were analyzed. The spectrograms of the output signals and the results of the objective SIGMOS metrics for the received topologies of signal transmission were obtained. During the analysis of the received spectrograms it was discovered that in scenarios with different noise environments, the Wiener filter works in a similar way to a low-pass filter, which can be partly used to suppress the high-frequency component or for restriction of the frequency range of the signal.

Better results were found in the case of using the RNNoise neural network in noise suppression tasks for typical scenarios with a noisy environment due to a fast response time to its change. According to the received statistical data, formed in table 1 and table 2, it was found that on average RNNoise shows the best results of coloration, noisiness and discontinuity.

**The scientific novelty** is that the method of adaptive filtering and noise suppression of the audio signal due to the predetermination of adaptive filtering coefficients in the process of machine learning based on neural networks has been further developed, which made it possible to improve the quality of noise suppression by an average of 20% for specific types of noise and increase the quality of noise suppression in different noise environments.

**The practical significance** of the study consists of running the RNNoise neural network on the ARM part of the hardware platform ZYNQ Zedboard using Vitis-IDE tools and obtaining objective signal metrics before and after processing using the SIGMOS package. Also, the hardware implementation of various adaptive filtering algorithms made it possible to prove the advantages of a filter based on the neural network for noise suppression in audio signals with different parameters.

**Prospects for further research** are related to the analysis of the possibility of the weight coefficients database formation of trained neural networks, which implement adaptive filtering for the corresponding types of noise. The computationally complex part of the neural network can be implemented using high-level synthesis and adapted for loading into the SoC. This will significantly reduce the noise suppression devices' design time and allow novice designers to use optimized and verified technical solutions for their specific tasks.

## ACKNOWLEDGEMENTS

The study was completed with the consultation support from the Logitech Audio R&D. The authors express their sincere gratitude.

## REFERENCES

1. Lahti S., Rintala M., Hamalainen T. D. Leveraging Modern C++ in High-level Synthesis, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2023, Vol. 42, № 4, pp. 1123–1132. DOI: 10.1109/TCAD.2022.3193646.
2. Monson J., Wirthlin M., Hutchings B. L. Optimization techniques for a high level synthesis implementation of the Sobel filter, *International Conference on reconfigurable computing and FPGAs (ReConFig'13)*. Cancun, Mexico, 9–11 December 2013, pp. 1–6. DOI: 10.1109/ReConFig.2013.6732315.
3. Plugariu O., Petrica L., Pirea R., Hobincu R. Hadoop ZedBoard cluster with GZIP compression FPGA acceleration, *11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI'19)*. Pitesti, Romania, 27–29 June 2019, pp. 1–5. DOI: 10.1109/ ecai46879.2019.9042006.
4. Rong X., Sun T., Zhang X., Hu Y., Zhu C. , Lu J. GTCRN: A speech enhancement model requiring ultralow computational resources, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2024)*. Seoul, Korea, 14–19 April 2024, pp. 971–975. DOI: 10.1109/icassp48485.2024.10448310.
5. Leiva L., Torrents-Barrena J., Vazquez M. FPGA-based accelerator for AI-toolbox reinforcement learning library, *IEEE Embedded Systems Letters*, 2023, Vol. 15, № 2, pp. 113–116. DOI: 10.1109/les.2022.3218168.
6. Fan H., Wang H., Che K., Wu Z. Design of FPGA deep neural network accelerator based on high-level synthesis, *5th International Academic Exchange Conference on Science and Technology Innovation (IAECST'23)*. Guangzhou, China, 8–10 December 2023, pp. 163–166. DOI: 10.1109/iaecst60924.2023.10502749.
7. Chang Y.-W., Huang C.-C., Hwang Y.-T. A face mask detection system based on high level synthesis and hardware software codesign, *IET International Conference on Engineering Technologies and Applications (IET-ICETA'22)*. Changhua, Taiwan, 14–16 October 2022, pp. 1–2. DOI: 10.1109/ieticeta56553.2022. 9971488.
8. Riazati M., Daneshtalab M., Sjödin M. , Lisper B. DeepFlexiHLS: Deep neural network flexible high-level synthesis directive generator, *IEEE Nordic Circuits and Systems Conference (NorCAS'22)*. Oslo, Norway, 25–26 October 2022, pp. 1–6. DOI: 10.1109/norcas 57515.2022.9934617.


DOI 10.15588/1607-3274-2024-4-16

9. Doumanidis C. C. , Anagnostou C. , Arvaniti E.-S., Papadopoulou A. RNNoise-Ex: hybrid speech enhancement system based on rnn and spectral features [Electronic resource], 2021, pp. 1–5. Access mode: https://arxiv.org/pdf/2105.11813. DOI: 10.48550/arXiv. 2105.11813
10. Cheng B., Zhang G., Tao X. , Wang S., Wu N., Chen M. An improved real-time noise suppression method based on RNN and long-term speech information, *3rd international symposium on automation, information and computing (ICSPCC'22)*. Beijing, China, China, 9–11 December 2022, pp. 476–481. DOI: 10.1109/ICASSP40776.2020.9054597.
11. Reddy C.K.A., Beyrami E., Pool J., Cutler R., Srinivasan S., Gehrke J. A scalable noisy speech dataset and online subjective test framework, *International Conference "Interspeech 2019"*. Graz, Austria, 15–19 September 2019. pp. 1816–1820. DOI: 10.21437/interspeech. 2019-3087.
12. Chen B. , Zhou Y., Ma Y., Liu H. A new real-time noise suppression algorithm for far-field speech communication based on recurrent neural network, *IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC'22)*. Xi'an, China, 17–19 August 2021, pp. 1–5. DOI: 10.1109/icspcc52875. 2021.9564530.
13. Zenodo: Python library for audio and music analysis Librosa [Electronic resource], 2024. Access mode: https://zenodo.org/records/11192913.
14. Scheibler R., Bezzam E., Dokmanić I. Pyroomacoustics: A Python package for audio room simulation and array processing algorithms, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'18*), Calgary. Canada, 15–20 April 2018, pp. 351–355. DOI: doi.org/10.1109/icassp.2018.8461310.
15. Upadhyay N., Jaiswal R. K. Single channel speech enhancement: using wiener filtering with recursive noise estimation, *Proceeding of the Seventh International Conference on Intelligent Human Computer Interaction*, 2016, Vol. 84, pp. 22–30. DOI: 10.1016/j.procs.2016.04.061.
16. Shkil A., Rahlis D., Filippenko I. , Kornijenko V., Rozhnova T. Automated design of embedded digital signal processing systems on SOC platform, *Innovative technologies and scientific solutions for industries*, 2024, No. 1 (27), pp. 192–203. DOI: https://doi.org/10.30837/ITSSI.2024.27.192
17. Naderi B., Cutler R. Subjective evaluation of noise suppression algorithms in crowdsourcing, *International Conference "Interspeech 2021"*. Brno, Czechia, 30 August –3 September 2021, pp. 2132–2136. DOI: 10.21437/interspeech.2021-343.
18. Catalin R. N., Saabas A., Cutler R., Naderi B. , Braun S., Branets S. Speech signal improvement challenge [Electronic resource], *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'24)*, 2024. Access mode: https://arxiv.org/pdf/2401.14444. DOI: 10.48550/arXiv.2401.14444.

Received 23.08.2024.
Accepted 22.10.2024.

УДК 681.326

## ВИКОРИСТАННЯ АДАПТИВНОЇ ФІЛЬТРАЦІЇ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ У СИСТЕМАХ ПРИДУШЕННЯ ШУМУ, РЕАЛІЗОВАНИХ НА ПЛАТФОРМІ SOC

**Шкіль О. С.** – канд. техн. наук, доцент кафедри автоматизації проєктування обчислювальної техніки, Харківський національний університет радіоелектроніки, Харків, Україна.

**Філіпенко О. І. –** канд. техн. наук, доцент кафедри інфокомунікаційної інженерії ім. В. В. Поповського, Харківський національний університет радіоелектроніки, Харків, Україна.

**Рахліс Д. Ю**. – канд. техн. наук, доцент кафедри автоматизації проєктування обчислювальної техніки, Харківський національний університет радіоелектроніки, Харків, Україна.

**Філіпенко І. В.** – канд. техн. наук, доцент кафедри автоматизації проєктування обчислювальної техніки, Харківський національний університет радіоелектроніки, Харків, Україна.

**Пархоменко А. В**. – канд. техн. наук, доцент кафедри програмних засобів, Національний університет «Запорізька політехніка», Запоріжжя, Україна.

**Корнієнко В. Р.** – аспірант кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Харків, Україна.

### АНОТАЦІЯ

**Актуальність.** Сучасні системи відео конференційного зв'язку працюють у різноманітному шумовому оточенні, тому актуальними завданнями є збереження чіткості мовлення та забезпечення швидкої адаптації до зміни цього оточення. При розробці вбудованих систем виникає необхідність знайти баланс між споживанням ресурсів, продуктивністю та якістю сигналу, отриманого після придушення шуму. Системи на кристалі дозволяють використовувати потужність як процесорних ядер, доступних на апаратній платформі, так і FPGA, для виконання складних обчислень, що сприяє підвищенню швидкодії або зменшенню навантаження на основні ядра SoC.

**Мета.** Проведення порівняльного аналізу якості придушення шуму у аудіо сигналах алгоритмом адаптивної фільтрації та алгоритмом фільтрації з використанням машинного навчання на основі нейронної мережі rnnoise в пристроях придушення шуму на технологічній платформі SoC.

**Метод.** Оцінка за допомогою об'єктивних метрик, аналіз спектрограм з використанням бібліотеки Librosa на Python. Навчання нейромережі та проєктування моделі виконується на основі інструментів Python та Torch. Для реалізації нейронної мережі на платформі SoC використовувався пакет Vitis IDE.

**Результати**. Виконано аналіз двох методів придушення шуму з використанням адаптивного фільтру Вейнера та нейронної мережі RNNoise. У розглянутих сценаріях було визначено що нейронна мережа показує кращі результати придушення шуму згідно до аналізу спектрограм та об'єктивних метрик.

**Висновки.** У роботі було виконано порівняльний аналіз ефективності алгоритмів придушення шуму на базі адаптивних фільтрів і нейронної мережі у сценаріях з різним шумовим оточенням. Були отримані результати об'єктивних метрик SIGMOS для оцінки якості отриманого аудіосигналу. Додатково була виконана перевірка можливості запуску нейронної мережі RNNoise на технологічній платформі SoC ZYNQ 7000.

**КЛЮЧОВІ СЛОВА:** вбудовані системи, системи на кристалі, FPGA, адаптивна фільтрація, алгоритми цифрової обробки сигналів, алгоритми придушення шуму, аудіо сигнали, машинне навчання, нейронні мережі.


DOI 10.15588/1607-3274-2024-4-16

**ЛІТЕРАТУРА**

1. Lahti S. Leveraging Modern C++ in High-level Synthesis / S. Lahti, M. Rintala, T. D. Hamalainen // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2023. – Vol. 42, № 4. – P. 1123–1132. DOI: 10.1109/TCAD.2022.3193646.
2. Monson J. Optimization techniques for a high level synthesis implementation of the Sobel filter / J. Monson, M. Wirthlin, B. L. Hutchings // International Conference on reconfigurable computing and FPGAs (ReConFig'13), Cancun, Mexico, 9–11 December 2013. – P. 1–6. DOI: 10.1109/ReConFig.2013.6732315.
3. Hadoop ZedBoard cluster with GZIP compression FPGA acceleration / [O. Plugariu, L. Petrica, R. Pirea, R. Hobincu] // 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI'19), Pitesti, Romania, 27–29 June 2019. – P. 1–5. DOI: 10.1109/ecai46879.2019.9042006.
4. GTCRN: A speech enhancement model requiring ultralow computational resources / [X. Rong, T. Sun, X. Zhang et al.] // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2024), Seoul, Korea, 14–19 April 2024. – P. 971–975. DOI: 10.1109/icassp48485.2024.10448310.
5. Leiva L. FPGA-based accelerator for AI-toolbox reinforcement learning library / L. Leiva, J. Torrents-Barrena, M. Vazquez // IEEE Embedded Systems Letters, 2023. – Vol. 15, № 2. – P. 113–116. DOI: 10.1109/les.2022.3218168.
6. Design of FPGA deep neural network accelerator based on high-level synthesis / [H. Fan, H. Wang, K. Che, Z. Wu] // 5th International Academic Exchange Conference on Science and Technology Innovation (IAECST'23), Guangzhou, China, 8–10 December 2023. – P. 163–166. DOI: 10.1109/iaecst60924.2023.10502749.
7. Chang Y.-W. A face mask detection system based on high level synthesis and hardware software codesign / Y.-W. Chang, C.-C. Huang, Y.-T. Hwang // IET International Conference on Engineering Technologies and Applications (IET-ICETA'22), Changhua, Taiwan, 14–16 October 2022. – P. 1–2. DOI: 10.1109/iet-iceta56553.2022.9971488.
8. DeepFlexiHLS: Deep neural network flexible high-level synthesis directive generator / [M. Riazati, M. Daneshtalab, M. Sjödin, B. Lisper] // IEEE Nordic Circuits and Systems Conference (NorCAS'22), Oslo, Norway, 25–26 October 2022. – P. 1–6. DOI: 10.1109/norcas57515.2022.9934617.
9. RNNoise-Ex: hybrid speech enhancement system based on rnn and spectral features [Electronic resource] / [C. C. Doumanidis, C. Anagnostou, E.-S. Arvaniti, A. Papadopoulou]. – 2021. – P. 1–5. – Access mode: https://arxiv.org/pdf/2105.11813. DOI: 10.48550/arXiv.2105.11813.
10. An improved real-time noise suppression method based on RNN and long-term speech information / [B. Cheng, G. Zhang, X. Tao et al.] // 3rd international symposium on automation, information and computing (ICSPCC'22), Beijing, China, China, 9–11 December 2022. – P. 476–481. DOI: 10.1109/ICASSP40776.2020.9054597.
11. A scalable noisy speech dataset and online subjective test framework / [C.K.A. Reddy, E. Beyrami, J. Pool et al.] // International Conference "Interspeech 2019", Graz, Austria, 15–19 September 2019. – P. 1816–1820. DOI: 10.21437/interspeech.2019-3087.
12. A new real-time noise suppression algorithm for far-field speech communication based on recurrent neural network / [B. Chen, Y. Zhou, Y. Ma, H. Liu] // IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC'22), Xi'an, China, 17–19 August 2021. – P. 1–5. DOI: 10.1109/icspcc52875.2021.9564530.
13. Zenodo: Python library for audio and music analysis Librosa [Electronic resource]. – 2024. – Access mode: https://zenodo.org/records/11192913.
14. Scheibler R. Pyroomacoustics: A Python package for audio room simulation and array processing algorithms / R. Scheibler, E. Bezzam, I. Dokmanić // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'18), Calgary, Canada, 15–20 April 2018. – P. 351–355. DOI: 10.1109/icassp.2018.8461310.
15. Upadhyay N. Single channel speech enhancement: using wiener filtering with recursive noise estimation / N. Upadhyay, R. K. Jaiswal // Proceeding of the Seventh International Conference on Intelligent Human Computer Interaction. – 2016. – Vol. 84. – P. 22–30. DOI: 10.1016/j.procs.2016.04.061.
16. Автоматизоване проєктування вбудованих систем цифрового оброблення сигналів на платформі SoC / [О. С. Шкіль, Д. Ю. Рахліс, І. В. Філіпенко та ін.] // Сучасний стан наукових досліджень та технологій в промисловості. – 2024. – No. 1 (27). – С. 72–83. DOI: https://doi.org/10.30837/ITSSI.2024.27.192
17. Naderi B. Subjective evaluation of noise suppression algorithms in crowdsourcing / B. Naderi, R. Cutler // International Conference "Interspeech 2021", Brno, Czechia, 30 August – 3 September 2021. – P. 2132–2136. DOI: 10.21437/interspeech.2021-343.
18. Speech signal improvement challenge [Electronic resource] / [R. N. Catalin, A. Saabas, R. Cutler et al.] // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'24). – 2024. – Access mode: https://arxiv.org/pdf/2401.14444. DOI: 10.48550/arXiv.2401.14444.


DOI 10.15588/1607-3274-2024-4-16

UDC 007.52:681.5.034:681.513.675:681.513.8

# CONSTRUCTING SENSOR SIGNAL PROCESSING CHANNEL FOR AUTONOMOUS ROBOTIC PLATFORMS

**Sytnikov V. S.** – Dr. Sc., Professor, Head of the Department of Computer Systems, Odessa Polytechnic National University, Odessa, Ukraine.

**Kudermetov R. K.** – PhD, Associate Professor, Head of the Department of Computer Systems and Networks, Zaporizhzhia Polytechnic National University, Zaporizhzhia, Ukraine.

**Stupen P. V.** – PhD, Associate Professor, Associate Professor of the Department of Computer Systems, Odessa Polytechnic National University, Odessa, Ukraine.

**Polska O. V.** – Senior lecturer of the Department of Computer Systems and Networks, National University Zaporizhzhia Polytechnic, Zaporizhzhia, Ukraine.

**Sytnikov T. V.** – Post-graduate student of the Department of the Computer Systems, Odessa Polytechnic National University, Odessa, Ukraine.

## ABSTRACT

**Context.** The development of autonomous mobile robotic platforms has advanced rapidly, especially in cyber-physical systems where integrating physical components and computational resources is vital. A key challenge in such platforms is the efficient real-time processing of sensor signals under limited computational resources, enabling robots to operate independently of human intervention. Traditional signal processing methods demand significant power, which may limit mobile platforms constrained by energy and resources. This research focuses on restructuring sensor signal processing channels using digital bandpass filters while overcoming technical challenges posed by limited resources.

**Objective.** The goal is to create an efficient method for processing sensor signals in autonomous mobile platforms with constrained resources. This involves using low-order bandpass filters, capable of adjusting their characteristics and improving quality through sequential connection of identical filters. Reducing the computational load allows for enhanced overall performance of cyber-physical systems, improving efficiency under changing conditions and enabling autonomous task completion. New computational formulas are also proposed to simplify the design and better utilize onboard resources.

**Method.** The improved method for constructing sensor signal processing channels uses identical low-order frequency-dependent components, sequentially connected to solve challenges faced by higher-order components. This approach simplifies coefficient calculations for cutoff frequencies and improves filter performance by increasing the order and quality. The method achieves a quasi-linear phase-frequency characteristic, ensuring minimal distortion in the processed signals, while significantly reducing computational requirements.

**Results.** The proposed method effectively reduces computational costs while maintaining high performance in sensor signal processing. The new formulas allow for calculating filter coefficients with fewer resources, suitable for autonomous systems. Modelling and experimental verification confirm that this method lowers the computational load and enhances filter performance, enabling more efficient sensor data processing, extended battery life, and improved system reliability.

**Conclusions.** This research presents an efficient approach to sensor signal processing for resource-constrained autonomous robotic platforms. Sequentially connecting identical frequency-dependent components reduces computational costs while maintaining high signal processing quality. These findings are recommended for real-time applications requiring efficient resource utilization, contributing to improved autonomy and adaptability in mobile robotic systems.

**KEYWORDS:** autonomous mobile robotic platform, frequency-dependent components, filters, frequency characteristics, cyber-physical system.

## ABBREVIATIONS

AGV is an automated guided vehicle;
AFC is an amplitude-frequency characteristic;
AMP is an autonomous mobile platform;
AMR is an autonomous mobile robot;
AMRP is an autonomous mobile robotic platform;
AP is an automated platform;
AU is an adjustment unit;
BPF is a bandpass filter;
CPS is a cyber-physical system;
EA is an evolutionary algorithm;
FDC is a frequency-dependent component;
FPGA is a field-programmable gate array;
HPF is a high pass filter;
LPF is a low pass filter;
MD is a multi-dimensional polynomial approximation;
NF is a notch filter;
OCM is an operator-controller module;
PFC is a phase-frequency characteristic;
SBC is a single-board computer.

## NOMENCLATURE

$a_0, a_1, a_2, b_1, b_2$ are the real coefficients of the filter transfer function numerator and denominator;

$c$ is a cutoff frequency level;

$D(\varpi)$ is a denominator of frequency response function;

$f$ is a linear frequency;

$f_0$ is a central frequency of the BPF's AFC;


DOI 10.15588/1607-3274-2024-4-17

$f_1$ is a left cutoff frequency of the BPF's AFC;

$f_2$ is a right cutoff frequency of the BPF's AFC;

$f_d$ is a sampling frequency;

$H(z)$ is a transfer function in z-domain;

$H(\varpi)$ is a frequency response function;

$H_n(\varpi)$ is a frequency response function of $n$ connected identical filters in series;

$N(\varpi)$ is a numerator of the frequency response function;

$n$ is a number of identical connected filters;

$Q$ is a quality factor of BPF;

$z$ is a variable in the complex plane;

$\varphi$ is a phase-frequency characteristic;

$\varpi$ is a normalized angular frequency;

$\varpi_{BP}$ is a bandwidth of the BPF;

$\varpi_c$ is an AFC cutoff frequency at level $c$;

$\varpi_p$ is a peak frequency of the AFC;

$\varpi_{1n}$ is a cutoff frequency at a new level $\sqrt[n]{c}$ by main AFC for quadratic equation;

$\varpi_{cn1}, \varpi_{cn2}$ is the cutoff frequencies at a new level $\sqrt[n]{c}$ by main AFC;

1L, 1R are the left and right cutoff frequencies at the level $c$;

2L, 2R are the left and right cutoff frequencies at the level $\sqrt[n]{c}$;

3L, 3R are projections of the left and right cutoff frequencies 2L and 2R onto the new AFC.

## INTRODUCTION

The modern development of society is characterized by the increasing use of robotic systems in production processes and society. Such systems include AMPs, mobile robots, and mobile manipulators, which are actively used in factories and enterprises to perform complex and monotonous tasks such as assembly, welding, painting, loading and unloading operations. In the social environment, they are employed for various service tasks: domestic, office, medical, monitoring, inspection, and search operations. It is especially important to note the role of AMPs in the military field, where they are used for both reusable and single-use specific tasks, as well as for reconnaissance, search, monitoring, escorting, and so on [1].

Within the framework of AMPs, it is possible to distinguish between AGV and AMR. APs are portable robots that use floor markings for movement and employ radio waves, video cameras and surveillance systems, magnets, and lasers for navigation. In the presence of obstacles, they typically stop. AMRPs, on the other hand, are platforms controlled by software based on signals from a large number of sensors and actuators, allowing them to identify their position and surroundings in space and move without operator intervention [2–5]. Moreover, modern AMRPs use elements of artificial intelligence to make decisions during task execution.

Recently, there has been intense development in CPS across various fields. To date, different definitions of such systems exist, but many authors agree that they represent a fusion of physical processes and cybernetic components. This combination enhances the intellectual capabilities of AMRPs in performing their functions [6–8]. In such systems, preliminary motion modelling is often used to ensure the safety of movement within a group of AMRPs and the completion of tasks [9, 10]. Reliable data on the state of the AMRP and its surrounding environment are essential for this purpose.

Typically, AMRPs are constrained by size, power supply, and computational capacity. This necessitates the development of software and hardware components with minimal energy consumption while ensuring task performance, as data processing is also limited by the implementation of processing algorithms. This fits well with modern concepts of AMRP development – Industry 4.0–6.0 [11, 12].

**The object of study** is the process of sensor signal processing in AMRPs without operator intervention. With a large number of sensors and the presence of interference, filtering is required, with the ability to adjust the parameters of the processing components in real-time. This need for adaptation to operating conditions enhances the reliability of the data for decision-making and task execution by the actuators.

**The subject of study** is the model and method for constructing the signal processing channel for sensors with the ability to adjust the characteristics and configuration to ensure the operational functionality of the path. Known methods [13–19] are complex, time-consuming, and costly, especially for single-use AMRPs.

**The purpose of the work** is to improve the efficiency of the FDCs of an AMRP, as a CPS, by reducing the computational complexity of their reconfiguration through the analysis and improvement of the sequential connection model of the components, taking into account the features of their frequency characteristics.

## 1 PROBLEM STATEMENT

Let's assume a CPS is given in the form of an AMRP interacting with natural objects. This platform consists of set of sensors, set of hardware components, set of software components, and set of actuators, all of which must operate in real-time to perform a given task. The primary constraints for such a platform are power supply limitations and computational resource limitations. The platform has both high-level and low-level subsystems, which prioritize the allocation of these resources.

For this platform, the task of constructing a sensor signal processing channel, which belongs to the low-level subsystems, can be represented as the problem of finding the components of the channel, determined after researching the possible structures and components of the channel. The components of the processing channel must meet the


DOI 10.15588/1607-3274-2024-4-17

requirements of simplicity, the ability for quick and minimal adjustments, and minimal costs for improving the quality of the channel while maintaining limited power consumption.

In this formulation, it is necessary to find formulas for calculating new cutoff frequencies for the channel component when $n$ identical low-order components are sequentially connected with a transfer function $H(z)$, in other words, $\varpi_{cn1,cn2} = f(a_i, b_i, n, c)$, and also show which phase distortions occur with such a connection $\varphi_n f(n)$.

## 2 REVIEW OF THE LITERATURE

Development a processing channel based on FDCs, several tasks are usually addressed: approximation, developing, and implementation. To create tunable components, various approaches and methods are used by the authors, allowing for control over the frequency characteristics of the components. Let's consider these approaches and methods using the example of digital filters, as standard control system links are described by similar transfer functions.

Many authors start the process by approximating and normalizing the transfer function. In [13], the authors propose transitioning to state space and adjusting characteristics there, using the Faust language to achieve the Chamberlin form. This approach is quite complex to implement in AMRP.

Another approach introduces additional parameters into the transfer function to control the cutoff frequency [14, 15]. Retuning is easily done on non-recursive filters since they do not have roots in the denominator of the transfer function. This idea of controlling the cutoff frequency then applied to recursive filters. Another approach involves using MD polynomial approximation for recursive filters with the introduction of additional parameters into the transfer function. These approaches require prior preparation, and with each retuning, the transfer function needs to be checked for stability.

It is known that digital non-recursive filters are always stable and have a linear PFC. However, since it is also necessary to change the filter order to improve its quality factor, higher-order filters are often required, which leads to increased delay, and this is a problem for real-time signal processing. Therefore, moving to recursive filters, the filter's numerator coefficients are approximated using polynomials with variable parameter. To solve this problem, the transfer function is transformed using semi-definite programming in the frequency domain [16].

The problem of developing a recursive filter has proven complex due to the presence of poles in its transfer function. This leads to a non-linear phase characteristic of the transfer function, and the amplitude characteristic is also shifted due to the quantization of the denominator polynomial coefficients, which leads to instability. In [17], numerous efforts were made to achieve an optimal filter characteristic using several optimization methods. A task was formulated and solved to develop a recursive filter with various constraints using gradient methods, which led to an optimal passband characteristic with an almost linear phase, and in some cases, absolute linearity was achieved. However, the obtained solutions were suboptimal in many cases due to the conversion of the multimodal filter design problem into convex optimization. To overcome suboptimality, the authors used EA. It should be noted that system identification was conducted in the time domain, while in the frequency domain, various error functions were developed for the obtained amplitude characteristics, bringing them closer to the desired one.

This approach resulted in an optimal recursive filter characteristic; however, the linearity of the phase characteristic was not improved. The EA approach was applied to lower-order recursive filters. This approach is the most suitable since the filter characteristic is stable, almost linear, and the amplitude characteristic is also accurate. However, there is a significant error at the edges of the bandpass. The retuning process is complex, and requires significant computational resources, though it provides a satisfactory characteristic, and stability must be checked during retuning.

The recursive filter structure has unique advantages and disadvantages. The main advantage of this filter structure is that it provides a frequency characteristic comparable to that of a higher-order non-recursive filter, resulting in fewer computations needed to implement the filter. However, recursive filters can be unstable due to the presence of feedback. As a result, they are more challenging to design, and caution is needed to prevent instability. Recursive filters have a non-linear phase characteristic, which may make them unsuitable for certain applications requiring a linear phase. Additionally, because recursive filters take into account past outputs due to feedback, they are typically more sensitive to quantization noise, making them harder to implement using 16-bit controllers or fixed-point microprocessors. Generally, 32-bit resources are preferable for implementing recursive filters.

For this reason, non-recursive filters are often recommended for microprocessor-based implementations due to their simplicity in description and realization [15, 16, 18]. However, many open questions remain regarding parameter retuning and increasing filter order.

When implementing AMRPs, several tasks need to be solved simultaneously. This requires developing a modular AMRP structure, often designed to solve current tasks and expand for future capabilities. Using multiple microprocessor blocks leads to a distributed system within a single AMRP.

For instance, in [19], a concept of a modular and distributed architecture for a robotic system is presented. The architecture is based on the OCM, which describes the adaptation of distributed OCM for AMRP, taking into account the requirements for such robots, including real-time constraints and safety. The presented architecture hierarchically divides the system into a three-layer structure of controllers and operators. Controllers directly in-


DOI 10.15588/1607-3274-2024-4-17

teract with all sensors and actuators in the system, requiring strict real-time constraints. However, the reflexive operator processes information from the controllers, which can be done using model-based principles with finite state machines. A cognitive operator is used to optimize the system.

Further development of AMRPs is connected to the advent of small SBCs as more energy-efficient and suitable solutions for mobile robot processor blocks. Multiple SBCs or microcontrollers are combined into a distributed system for greater computational power or better reliability, and modular solutions make it easy to expand AMRP capabilities. This approach positions AMRP as a CPS capable of making decisions to improve signal processing quality in the processing path under uncertainty, without operator involvement.

At present, there are many definitions of CPS. In general, CPS are systems consisting of various natural objects, artificial subsystems, and control controllers, allowing such a formation to be considered a unified whole. CPS provide close coordination between computational and physical resources. Computers monitor and control physical processes using feedback loops, where events in the physical systems influence the computations and vice versa. The main technological prerequisites for the emergence of CPS are [20–24]:

– growth in the number of devices with embedded microcontrollers, microprocessors and data storage;

– integration that achieves maximum efficiency by combining individual components into larger systems;

– availability of numerous sensors and actuators;

– limitations in human cognitive abilities, which evolve more slowly than machines, leading to a point where they can no longer handle the volume of information required for decision-making.

Typically, this is an embedded system, a special-purpose system where the computational element is entirely integrated into the device it controls. Unlike a general-purpose computer, an embedded system performs one or several predefined tasks, usually with very specific real-time requirements.

From a technical perspective, an embedded system interacts with its environment in a controlled manner, meeting several requirements for ensuring the quality and timeliness of information necessary for task control and execution.

CPS integrate cybernetics, hardware and software technologies, and qualitatively new actuators embedded in their environment, capable of sensing changes, reacting to them, self-learning, and adapting.

It is worth noting that the general CPS architecture is divided into five fundamental levels [20–22]:

– physical level (lays the foundation for CPS architecture);

– network level (packet routing based on transforming a unique identifier assigned to each active device into network-based identifiers);

– transport level (packets are fragmented into smaller parts);

– intermediate level (terminal management, protocol conversion);

– application level (stores, analyzes, and updates information).

Thus, the development of the sensor signal processing channel fits into this CPS architecture, interacting actively with sensors to provide the decision-making system and actuators with reliable information. The path must be able to adjust its properties according to the operational conditions of the AMRP.

## 3 MATERIALS AND METHODS

In building the sensor signal processing path for AMRP, there are two main tasks: adjusting the cutoff frequency or bandpass and increasing the quality factor of the AFC or enhancing the slope of the AFC rise and fall.

In addressing the first task, adjusting the cutoff frequency for LPF and HPF, as well as the centre frequency and bandwidth (rejection) for bandpass and NF, an analysis of methods for adjusting recursive filters of various orders was conducted. Several methods for tuning both analog and digital filters were considered, some of which are reflected in the works [13–19]. Particular attention was paid to the simplicity of the method, the computational load for recalculating new filter coefficients, stability, and the interdependencies between filter coefficients. It should be noted that as the filter order increases, the interdependence of coefficients during re-tuning grows. This leads to a complex recalculation process, the range limitations, and the need for mandatory stability checks. For digital filters, the quantization of coefficients also plays a significant role. Based on this analysis, the low-order digital filters (first and second order) and an analytical method for calculating new coefficients are recommended [25].

To utilize the analytical method, models of FDCs were developed, allowing the creation of systems of equations from which calculation formulas for retuning were derived. The limited number of coefficients simplified the calculations and reduced computational costs, and the stability check was found to be fairly straightforward.

Enhancing the “intellectual capabilities” of the channel during retuning is possible based on an adaptive approach, as shown in [14]. The peculiarity of controlling the main filter lies in the AU, which also defines its intelligent properties. Currently, AUs widely use comb filters for analyzing the input and output signals of the main filter.

Based on the obtained theoretical and practical data, the task of improving the quality factor of the AFC or increasing the steepness of the rise and fall of the AFC is proposed to be solved using similar FDCs of a low order.

It is known that when components are connected in series, the resulting transfer function is the product of the individual transfer functions of the components. If these components are of the same type, this multiplication turns into raising the individual transfer function to a power. Moving to the complex transfer coefficient based on complex variable theory, it is possible to point out that the


DOI 10.15588/1607-3274-2024-4-17

module, i.e., AFC is raised to the corresponding power, and the phase, i.e., PFC is multiplied by the indicator of the degree.

It is also known that the quality factor of BPFs is determined by the AFC. To do this, the bandwidth is defined as the difference between the cutoff frequencies at a level of 0.707 (approximately 3 dB) from the peak of the AFC, as half of the signal's power is concentrated at the filter's output at this level. Thus, the quality factor $Q$ is determined as follows

$$Q = \frac{f_0}{f_2 - f_1}. \quad (1)$$

Naturally, the higher the quality factor, the steeper the AFC, i.e., the rate of rise and fall of the characteristic. Let's consider such a configuration using first-order identical BPFs as an example since they are often used in AMRP. When connecting such filters in series, the AFC of the new configuration becomes "compressed", with the cutoff frequencies shifting toward the central frequency, and the AFC slope increases, thus improving the quality factor of this configuration (Fig. 1) [26].

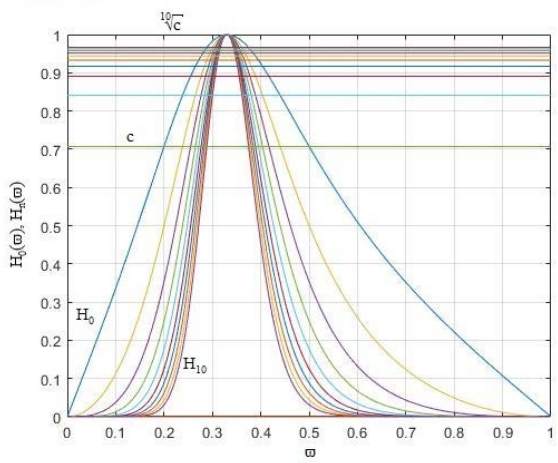



Figure 1 – AFC of first-order Butterworth digital BPFs connected in series

The transfer function of the main first-order BPF in $z$-domain is described as follows

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}}. \quad (2)$$

For $z = e^{j\varpi}$, $a_1 = 0$ and $a_2 = -a_0$, the transfer function (2) can be transformed into the following frequency response function

$$H(\varpi) = \frac{N(\varpi)}{D(\varpi)}, \quad (3)$$

where

$$N(\varpi) = 2a_0 \sin(\varpi),$$

$$D(\varpi) = \sqrt{(1-b_2)^2 + b_1^2 + 2b_1(1+b_2)\cos(\varpi) + 4b_2\cos^2(\varpi)}.$$

It should be noted that the peak frequency of the AFC remains unchanged and is determined by the equation

$$\varpi_p = \arccos\left(-\frac{b_1}{1+b_2}\right).$$

Usually, the cutoff frequency level is defined as $c = 0.707$, then solving equation $H(\varpi_c) = c$, we determine the two cutoff frequencies 1L and 1R (Fig. 2).

When multiplying identical AFCs or raising them to a power, the level remains the same, but to determine the cutoff frequencies of the AFC of the new configuration, when connected in series, it is necessary to extract the corresponding root from the level $c$, i.e., $\sqrt[n]{c}$ (Fig. 2). In Fig. 1, these levels are shown as horizontal lines.

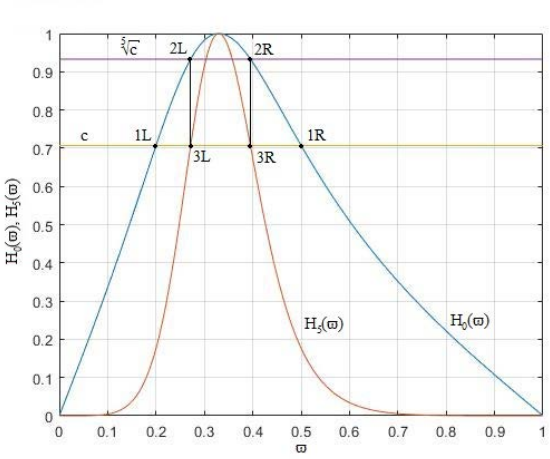



Figure 2 – AFC of the main first-order Butterworth filter and AFC when connecting five identical first-order Butterworth filters in series

In this case, from the AFC of the main filter, it is possible to calculate the cutoff frequencies of the AFC of the new configuration. Figure 2 shows the correspondence between the cutoff frequencies of the main first-order AFC at the level $c$ and the AFC when five identical first-order AFCs are connected in series. These cutoff frequencies are determined by the main AFC, whose parameters are known, and by the new level equal to $\sqrt[n]{c}$. Figure 2 shows the level $c$ for the main filter and its cutoff frequencies 1L and 1R, then the new level is determined from the number of connections $\sqrt[n]{c}$ and based on it, the new cutoff frequencies 2L and 2R are determined. Based on these, as a projection onto the new AFC of the connection, it can be possible to get 3L and 3R [27].


DOI 10.15588/1607-3274-2024-4-17

To determine the cutoff frequencies of the new AFC after connecting $n$ identical filters based on the main AFC, it is necessary to solve the equation (4):

$$H^2(\varpi) = \frac{N^2(\varpi)}{D^2(\varpi)} = \sqrt[n]{c^2}. \quad (4)$$

The substitutions $a_0 = (1-b_2)/2$ and $\sin(\varpi) = \sqrt{1-\cos^2(\varpi)}$ to the numerator of (3) yield the following expression for $N^2(\varpi)$:

$$N^2(\varpi) = (1-b_2)^2(1-\cos^2(\varpi)).$$

After transforming equation (4) can be written in the form of a quadratic equation for $\cos(\varpi)$:

$$\begin{aligned}[4b_2c^{\frac{2}{n}} + (1-b_2)^2]\cos^2(\varpi) + 2b_1(1+b_2)\cos(\varpi) + \\ + c^{\frac{2}{n}}b_1^2 + (1-b_2)^2(c^{\frac{2}{n}} - 1) = 0.\end{aligned} \quad (5)$$

Solving this equation gives formulas for determining the cutoff frequencies when $n$ identical filters are connected in series:

$$\varpi_{cn1,cn2} = \arccos\left( \frac{-b_1(1+b_2)c^{\frac{2}{n}}}{4b_2c^{\frac{2}{n}} + (1-b_2)^2} \pm \right.$$
$$\left. \pm \frac{(1-b_2)\sqrt{\left(1-c^{\frac{2}{n}}\right)\left[\left(4b_2 - b_1^2\right)c^{\frac{2}{n}} + (1-b_2)^2\right]}}{4b_2c^{\frac{2}{n}} + (1-b_2)^2} \right). \quad (6)$$

When creating such a new signal processing path for sensors, a question arises as to how the PFC will look compared to the direct calculation of a filter of the given order and how linear it will be. The research of changes in the PFC was based on the transfer function of the main filter (2). For this, an analytical mathematical description of the PFC was obtained, which after transformation looks as follows

$$\varphi = \arctan\left( \frac{b_1 + (b_2+1)\cos(\varpi)}{(1-b_2)\sin(\varpi)} \right). \quad (7)$$

As mentioned earlier, when connecting $n$ identical filters in series, the PFC of the new configuration is the multiplication of the PFC of the main filter and the number of main filters connected, i.e.,

$$\varphi_n = n \cdot \varphi. \quad (8)$$

For other filters, when similar components are connected in series, the AFC changes as follows:

– for the LPF, it shifts to the left, i.e., towards lower frequencies with the quality factor increases;

– for the HPF, the characteristic shifts to the right, i.e., towards higher frequencies with the quality factor increases;

– for the NF, unlike BPF, the characteristic does not shift toward the central frequency, meaning it “expands” outward from the central frequency rather than “compressing”.

Such AFC behavior is not always suitable for solving the signal processing tasks required for sensor data.

## 4 EXPERIMENTS

As a result of theoretical research, corresponding calculations, modelling, and experimental verifications were carried out using real AMRP.

In solving the first task of retuning, the computational costs for recalculating the transfer function coefficients were assessed. The computational load was evaluated by the number of operations required to implement the proposed formulas. The results were compared with the method based on variable substitution. Each processor has its own values for the execution of addition operations. Raspberry Pi 4 was chosen for evaluation. The results were experimentally verified using it. Based on the results obtained, the proposed method of recalculating transfer function coefficients is 27% more efficient for first-order LPF and HPF, 29% for second-order LPF and HPF, and 13% for BPF and NF.

Modelling and experimental verification of serially connected similar BPFs resulted in a quasi-linear PFC compared to equivalent filters calculated directly.

## 5 RESULTS

When verifying the obtained relationships, for example, Fig. 3 shows the cutoff frequency dependencies derived from formula (7).

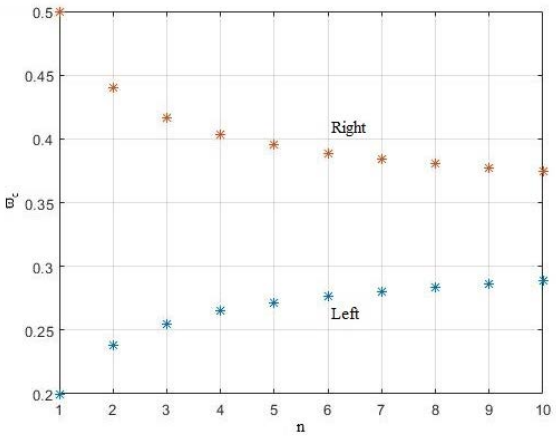



Figure 3 – Graph of the cutoff frequency dependence on the number of connected identical filters when filters connected in series


DOI 10.15588/1607-3274-2024-4-17

Based on these ratios, the bandwidth of such a connection is determined as $\varpi_{BP} = |\varpi_{cn1} - \varpi_{cn2}|$ (Fig. 4).

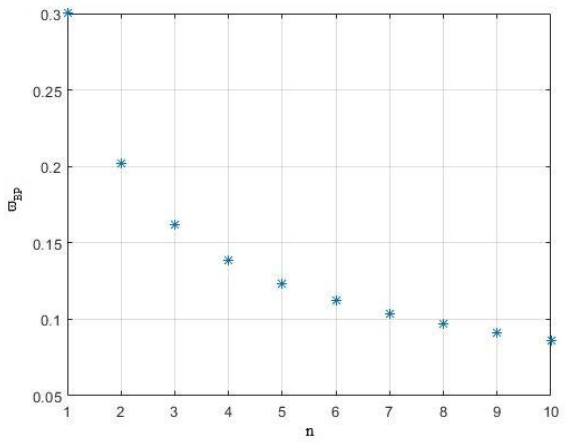



Figure 4 – Graph of bandwidth $\varpi_{BP}$ dependence on the number of connections

As seen in Fig. 4, the bandwidth decreases exponentially. At the same time, it can be shown how much the bandwidth decreases with a series connection, and accordingly, the quality factor of such a configuration increases in accordance with (1) (Fig. 5).

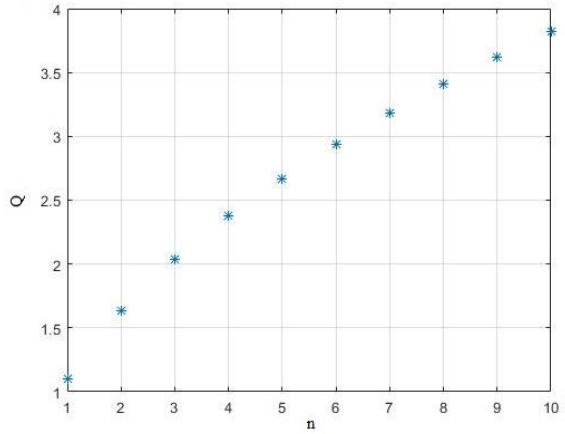



Figure 5 – Graph of the quality factor Q dependence on the number of connections

For example, with connecting four identical components, the bandwidth decreases by more than twice, with connecting eight – by more than three times, with ten – the bandwidth is reduced by 3.5 times. At the same time, the quality factor increases by 2.4 times, 3.4 times, and 3.8 times, respectively.

Modelling and experimental verification of the series connection of identical BPFs lead to a quasi-linear PFC compared to similar directly calculated filters (Fig. 6).

Figure 6 shows that the PFC of the new configuration, according to equation (8), i.e., $\varphi_5 = 5 \cdot \varphi$, five identical filters, is quasi-linear compared to the PFC of a fifth-order filter calculated directly. Additionally, when calculating an identical filter, only three coefficients are computed, whereas the fifth-order filter requires six coefficients in the numerator and ten in the denominator, totaling 16 coefficients that need recalculating and verification for stability. For filters of this order, this is a challenging task. Verifying the stability of first-order BPFs is a well-known and straightforward task. Moreover, the serial connection of stable filters results in the stability of the entire configuration.

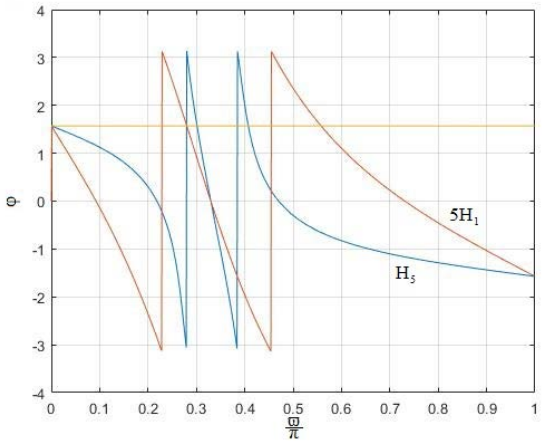



Figure 6 – Graph comparing the PFC of a series connection of five identical filters with the transfer function (1) ($5H_1$ is the transfer function after connecting five main filters) and the PFC of a fifth-order directly calculated filter $H_5$

When analyzing the PFC of the serial connection of identical components, it was found that the PFC of the new configuration is significantly more linear compared to the directly calculated PFC of the required order. Under limited computational capabilities, the proposed approach is considerably better than the traditional one.

## 6 DISCUSSION

The technical implementation of this approach in digital form can be both hardware and software-based. The hardware implementation relies on the serial connection of n identical components. The main component is calculated based on data stream information and operating modes. Cutoff frequencies, bandwidth, and gain coefficients are determined from this.

In the hardware implementation, one option considered was the serial connection of multiple identical filters with registers at the outputs of these components. A commutator handles the commutation of register outputs to the device output, which reduces filter switching time and the transition process time since the necessary data is already in the registers. This implementation could also be realized on an FPGA. However, this approach increases the energy consumption of the entire processing pipeline.

The software implementation was applied for processing data from the ultrasonic obstacle sensor on an AMRP using identical first-order components. This solution was convenient for both implementation and operation, reducing computation time since some constants were precomputed and stored in memory cells. Additionally, there were written a program for increasing filter order and the steepness of the AFC. However, the transition process time increased.


DOI 10.15588/1607-3274-2024-4-17

## CONCLUSIONS

A relevant problem has been solved: the comprehensive retuning of sensor signal processing characteristics with limited computational capabilities onboard AMRP without operator involvement. This was demonstrated using digital BPFs.

**The scientific novelty** of the results lies in the fact that, based on a model of serially connected identical FDCs described by first- and second-order transfer functions with specific frequency characteristics, new calculation formulas for such a connection were obtained for the first time. These formulas reduce computational costs when calculating coefficients based on given cutoff frequencies. Moreover, this connection allows for an increase in the order and quality factor of the new configuration, as well as achieving a quasi-linear PFC.

**The practical significance** of the results is that the obtained relationships enable the calculation of new cutoff frequency values with lower computational costs. Modelling and experimental verification results recommend this method and its formulas for practical use in AMRP with limited real-time computational resources.

**Prospects for further research** include extending this method to a wide range of practical tasks in robotics.

## ACKNOWLEDGEMENTS

The work was supported by the state budget research project of the Odessa Polytechnic National University "Models, methods of development and hardware solutions for high-speed components of mobile programmable systems" (state registration number 0121U110245), as well as by contract works No. 1880 and No. 1922.

## REFERENCES

1. Understanding AMR Robots: A Comprehensive Guide [Electronic resource]. Access mode: https://www.wevolver.com/article/understanding-amr-robots-a-comprehensive-guide/
2. What do AGV and AMR stand for? [Electronic resource]. – Access mode: https://mobile-industrial-robots.com/blog/agv-vs-amr-whats-the-difference/
3. Sensors in Autonomous Mobile Robots for Localization and Navigation [Electronic resource]. Access mode: https://www.techtarget.com/iotagenda/blog/IoT-Agenda/Sensors-in-autonomous-mobile-robots-for-localization-and-navigation/
4. What are the Three Types of Robotic Systems [Electronic resource]. Access mode: https://www.international-development.eu/what-are-the-three-types-of-robotic-systems/
5. What is the Difference Between Automation and Robotics [Electronic resource]. Access mode: https://robotnik.eu/what-is-the-difference-between-automation-and-robotics/
6. Puente-Castro A., Rivero D., Pazos A. et al. A Review of Artificial Intelligence Applied to Path Planning in UAV Swarms, *Neural Computing and Applications*, 2021, Vol. 34, No. 1. – P. 153–170. DOI: 10.1007/s00521-021-06569-4
7. Obukhova K., Zhuravska I., Burenko V. Diagnostics of Power Consumption of a Mobile Device Multi-core Processor with Detail of Each Core Utilization, *Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) : 15th IEEE International Conference*, Lviv-Slavske, 25–29 February 2020 : proceedings, IEEE, 2020, pp. 368–372. DOI: 10.1109/TCSET49122.2020.235456
8. Obukhova K., Savchuk T., Zhuravska I. et al. Prevention of Unmanned Vessels Collisions Due to Pre-modeling the Remote Control Center CPU Load, *Computer Sciences and Information Technologies (CSIT) : 16th IEEE International Conference*, Lviv, 22–25 September 2021 : proceedings, IEEE, 2021, Vol. 2, pp. 202–205. DOI: 10.1109/CSIT52700.2021.9648800
9. Semenov S. Voloshyn D., Ahmed A. N. Mathematical Model of the Implementation Process of Flight Task of Unmanned Aerial Vehicle in the Conditions of External Impact, *International Journal of Advanced Trends in Computer Science and Engineering*, 2019, Vol. 8, No. 1, pp. 7–13. DOI: 10.30534/ijatcse/2019/0281.22019
10. Voloshyn D., Brechko V., Semenov S. Method of an Unmanned Aerial Vehicle Composition Route in Space, *Advanced Information Systems*, 2021, Vol. 5, No. 4, pp. 26–33. DOI: 10.20998/2522-9052.2021.4.04
11. Industry 4.0: An Overview [Electronic resource]. Access mode: https://www.researchgate.net/publication/326352993_Industry_40_an_overview
12. Adel A. Future of Industry 5.0 in Society: Human-Centric Solutions, Challenges and Prospective Research Areas, *Journal of Cloud Computing*, 2022, Vol. 11, Article ID 40. DOI: 10.1186/s13677-022-00314-5
13. Digital State-Variable Filters [Electronic resource]. Access mode: https://ccrma.stanford.edu/~jos/svf/svf.pdf
14. Koshita S. Abe M., Kawamata M.Recent Advances in Variable Digital Filters. In book: Digital Systems, [Electronic resource]. Access mode: https://www.researchgate.net/publication/329258883_Recent_Advances_in_Variable_Digital_Filters
15. Dhabu S., Ambede A., Agrawal N. et al. Variable Cutoff Frequency FIR Filters: A Survey, *SN Applied Sciences*, 2020, Vol. 2, Article ID 343. DOI: 10.1007/s42452-020-2140-6
16. Ubayama N., Miyata T., Aikawa N. A Design Method for IIR Digital Filters with Variable Stopband Using Semidefinite Programming, *Information Theory & Its Applications : IEEE International Symposium*. Taichung, Taiwan, 17–20 October 2010 : proceedings, IEEE, 2010, pp. 611–615. DOI: 10.1109/ISITA.2010.5649706
17. Agrawal N. Kumar A., Bajaj V. et al. Design of Digital IIR Filter: A Research Survey, *Applied Acoustics*, 2021, Vol. 172, Article ID 107669. DOI: 10.1016/j.apacoust.2020.107669
18. Using Microcontrollers in Digital Signal Processing Applications [Electronic resource]. Access mode: https://www.silabs.com/documents/public/application-notes/an219.pdf
19. Jahn U., Wolff C., Schulz P. Concepts of a Modular System Architecture for Distributed Robotic Systems, *Computers*, 2019, Vol. 8(1), Article ID 25. DOI: 10.3390/computers8010025
20. Jianjun S., Xu W., Jizhen G. et al. The Analysis of Traffic Control Cyber-physical Systems, *Procedia-Social and Behavioral Sciences*, 2013, Vol. 96, pp. 2487–2496. DOI: 10.1016/j.sbspro.2013.08.278
21. Kaur U. 463 Cyber-Physical Systems with robots and AI for precision dairy farming, *Journal of Animal Science*, 2024, Vol. 102, pp. 297–298. DOI: 10.1093/jas/skae234.340


DOI 10.15588/1607-3274-2024-4-17

22. Burns A., Hayes I. J., Jones C. B.Deriving Specifications of Control Programs for Cyber Physical Systems, *The Computer Journal*, 2020, Vol. 63, No. 5, pp. 774–790. DOI: 10.1093/comjnl/bxz019
23. Halder S., Afsari K., Akanmu A. A Robotic Cyber-Physical System for Automated Reality Capture and Visualization in Construction Progress Monitoring, [Electronic resource]. Access mode: https://10.48550/arXiv.2402.07034
24. Mañas-Álvarez F.J., Guinaldo M., Dormido R. et al. Scalability of Cyber-Physical Systems with Real and Virtual Robots in ROS 2, *Sensors*, 2023, Vol. 23, Article ID 6073. DOI: 10.3390/s23136073
25. Ukhina H., Sytnikov V., Streltsov O. et al. Application of the Computer System Component with Adjustment Elements for Processing Sensor Signals, *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) : 11th IEEE International Conference.* Cracow, 22–25 September 2021 : proceedings, IEEE, 2021, Vol. 1, pp. 12–17. DOI: 10.1109/IDAACS53288.2021.9660992
26. Quality Factor and Bandwidth [Electronic resource]. Access mode: https://ecstudiosystems.com/discover/textbooks/ basic-electronics/filters/quality-factor-and-bandwidth/
27. Sytnikov T., Sytnikov V., Streltsov O. et al.] /Increasing Order of Digital Filters of the Same Type in Robotic Systems, *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) : 12th IEEE International Conference*. Dortmund, 07–09 September 2023, proceedings, IEEE, 2023, Vol. 1, P. 1200–1204. DOI: 10.1109/IDAACS58523.2023

Received 09.08.2024.
Accepted 23.10.2024.

УДК 007.52:681.5.034:681.513.675:681.513.8

## ПОБУДОВА КАНАЛУ ОБРОБКИ СИГНАЛІВ ДАТЧИКІВ ДЛЯ АВТОНОМНОЇ РОБОТОТЕХНІЧНОЇ ПЛАТФОРМИ

**Ситніков В. С.** – д-р техн. наук, професор, завідувач кафедри комп'ютерних систем Національного університету «Одеська політехніка», Одесса, Україна.

**Кудерметов Р. К.** – канд. техн. наук, доцент, завідувач кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка», Запоріжжя, Україна.

**Ступень П. В.** – канд. техн. наук, доцент кафедри комп'ютерних систем Національного університету «Одеська політехніка», Одесса, Україна.

**Польська О. В.** – старший викладач кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка», Запоріжжя, Україна.

**Ситніков Т. В.** – аспірант кафедри комп'ютерних систем Національного університету «Одеська політехніка», Одесса, Україна.

### АНОТАЦІЯ

**Актуальність.** Розвиток автономних мобільних роботизованих платформ швидко прогресує, особливо в області кіберфізичних систем, де важлива інтеграція фізичних компонентів і обчислювальних ресурсів. Одним із ключових викликів для таких платформ є ефективна обробка сигналів датчиків у режимі реального часу за умов обмежених обчислювальних ресурсів, що дозволяє роботам діяти незалежно від втручання людини. Традиційні методи обробки сигналів вимагають значних ресурсів, що може стати проблемою для платформ з обмеженою енергією та ресурсами. Це дослідження зосереджене на перебудові каналу обробки сигналів за допомогою цифрових смугових фільтрів, долаючи технічні труднощі, що виникають через обмеження ресурсів.

**Мета роботи** – створення ефективного методу обробки сигналів датчиків на автономних мобільних платформах з обмеженими ресурсами. Це включає використання низько-порядкових смугових фільтрів, які можуть змінювати свої характеристики і підвищувати якість за допомогою послідовного з'єднання однакових фільтрів. Зниження обчислювального навантаження покращує загальну продуктивність кіберфізичних систем, підвищуючи ефективність роботи в умовах змін та дозволяючи автономне виконання завдань. Запропоновані нові розрахункові формули спрощують процес проєктування фільтрів та дозволяють ефективніше використовувати обмежені ресурси платформи.

**Метод.** Покращений метод побудови каналів обробки сигналів використовує однакові низько-порядкові частотно-залежні компоненти, послідовно з'єднані для вирішення проблем, характерних для високопорядкових компонентів. Такий підхід спрощує обчислення коефіцієнтів для заданих частот зрізу та підвищує продуктивність фільтра завдяки збільшенню порядку та якості. Метод досягає квазілінійної фазо-частотної характеристики, що мінімізує спотворення сигналу, і значно знижує обчислювальні вимоги.

**Результати.** Запропонований метод ефективно знижує обчислювальні витрати при збереженні високої продуктивності в обробці сигналів датчиків. Нові формули дозволяють розраховувати коефіцієнти фільтрів з використанням меншої кількості ресурсів, що робить їх придатними для автономних систем. Моделювання та експериментальна перевірка підтверджують, що цей метод знижує навантаження та покращує частотні характеристики фільтрів, дозволяючи роботам більш ефективно взаємодіяти з оточенням у режимі реального часу. Підвищена ефективність обробки сигналів також подовжує час роботи та підвищує надійність системи.

**Висновки.** Це дослідження пропонує ефективний підхід до обробки сигналів для автономних мобільних платформ з обмеженими ресурсами. Метод послідовного з'єднання однакових частотно-залежних компонентів знижує обчислювальні витрати та підтримує високу якість обробки сигналів. Результати моделювання та експериментів підтверджують ефективність нових розрахункових формул для покращення продуктивності системи. Цей підхід добре підходить для кіберфізичних систем, де критично важлива робота в реальному часі та ефективне використання ресурсів.

**КЛЮЧОВІ СЛОВА:** автономна мобільна робототехнічна платформа, частотно-залежні компоненти, фільтри, частотні характеристики, кіберфізична система.


DOI 10.15588/1607-3274-2024-4-17

**ЛІТЕРАТУРА**

1. Understanding AMR Robots: A Comprehensive Guide [Electronic resource]. – Access mode: https://www.wevolver.com/article/understanding-amr-robots-a-comprehensive-guide/
2. What do AGV and AMR stand for? [Electronic resource]. – Access mode: https://mobile-industrial-robots.com/blog/agv-vs-amr-whats-the-difference/
3. Sensors in Autonomous Mobile Robots for Localization and Navigation [Electronic resource]. – Access mode: https://www.techtarget.com/iotagenda/blog/IoT-Agenda/Sensors-in-autonomous-mobile-robots-for-localization-and-navigation/
4. What are the Three Types of Robotic Systems [Electronic resource]. – Access mode: https://www.international-development.eu/what-are-the-three-types-of-robotic-systems/
5. What is the Difference Between Automation and Robotics [Electronic resource]. – Access mode: https://robotnik.eu/what-is-the-difference-between-automation-and-robotics/
6. A Review of Artificial Intelligence Applied to Path Planning in UAV Swarms / [A. Puente-Castro, D. Rivero, A. Pazos et al.] // Neural Computing and Applications. – 2021. – Vol. 34, No. 1. – P. 153–170. DOI: 10.1007/s00521-021-06569-4
7. Obukhova K. Diagnostics of Power Consumption of a Mobile Device Multi-core Processor with Detail of Each Core Utilization / K. Obukhova, I. Zhuravska, V. Burenko // Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) : 15th IEEE International Conference, Lviv-Slavske, 25–29 February 2020 : proceedings. – IEEE, 2020. – P. 368–372. DOI: 10.1109/TCSET49122.2020.235456
8. Prevention of Unmanned Vessels Collisions Due to Premodeling the Remote Control Center CPU Load. / [K. Obukhova, T. Savchuk, I. Zhuravska et al.] // Computer Sciences and Information Technologies (CSIT) : 16th IEEE International Conference, Lviv, 22–25 September 2021 : proceedings. – IEEE, 2021. – Vol. 2. – P. 202–205. DOI: 10.1109/CSIT52700.2021.9648800
9. Semenov S. Mathematical Model of the Implementation Process of Flight Task of Unmanned Aerial Vehicle in the Conditions of External Impact / S. Semenov, D. Voloshyn, A. N. Ahmed // International Journal of Advanced Trends in Computer Science and Engineering. – 2019. – Vol. 8, No. 1. – P. 7–13. DOI: 10.30534/ijatcse/2019/0281.22019
10. Voloshyn D. Method of an Unmanned Aerial Vehicle Composition Route in Space / D. Voloshyn, V. Brechko, S. Semenov // Advanced Information Systems. – 2021. – Vol. 5, No. 4. – P. 26–33. DOI: 10.20998/2522-9052.2021.4.04
11. Industry 4.0: An Overview [Electronic resource]. – Access mode: https://www.researchgate.net/publication/326352993_Industry_40_an_overview
12. Adel A. Future of Industry 5.0 in Society: Human-Centric Solutions, Challenges and Prospective Research Areas / A. Adel // Journal of Cloud Computing. – 2022. – Vol. 11, Article ID 40. DOI: 10.1186/s13677-022-00314-5
13. Digital State-Variable Filters [Electronic resource]. – Access mode: https://ccrma.stanford.edu/~jos/svf/svf.pdf
14. Koshita S. Recent Advances in Variable Digital Filters. In book: Digital Systems / S. Koshita, M. Abe, M. Kawamata [Electronic resource]. – Access mode: https://www.researchgate.net/publication/329258883_Recent_Advances_in_Variable_Digital_Filters
15. Variable Cutoff Frequency FIR Filters: A Survey / [S. Dhabu, A. Ambede, N. Agrawal et al.] // SN Applied Sciences. – 2020. – Vol. 2, Article ID 343. DOI: 10.1007/s42452-020-2140-6
16. Ubayama N. A Design Method for IIR Digital Filters with Variable Stopband Using Semidefinite Programming / N. Ubayama, T. Miyata, N. Aikawa // Information Theory & Its Applications : IEEE International Symposium, Taichung, Taiwan, 17–20 October 2010 : proceedings. – IEEE, 2010. – P. 611–615. DOI: 10.1109/ISITA.2010.5649706
17. Agrawal N. Design of Digital IIR Filter: A Research Survey / [N. Agrawal, A. Kumar, V. Bajaj et al.] // Applied Acoustics. – 2021. – Vol. 172, Article ID 107669. DOI: 10.1016/j.apacoust.2020.107669
18. Using Microcontrollers in Digital Signal Processing Applications [Electronic resource]. – Access mode: https://www.silabs.com/documents/public/application-notes/an219.pdf
19. Jahn U. Concepts of a Modular System Architecture for Distributed Robotic Systems / U. Jahn, C. Wolff, P. Schulz // Computers. – 2019. – Vol. 8(1), Article ID 25. DOI: 10.3390/computers8010025
20. The Analysis of Traffic Control Cyber-physical Systems / [S. Jianjun, W. Xu, G. Jizhen et al.] // Procedia-Social and Behavioral Sciences. – 2013, Vol. 96. – P. 2487–2496. DOI: 10.1016/j.sbspro.2013.08.278
21. Kaur U. 463 Cyber-Physical Systems with robots and AI for precision dairy farming / U. Kaur // Journal of Animal Science. – 2024. – Vol. 102. – P. 297–298. DOI: 10.1093/jas/skae234.340
22. Burns A. Deriving Specifications of Control Programs for Cyber Physical Systems / A. Burns, I. J. Hayes, C. B. Jones // The Computer Journal. – 2020. – Vol. 63, No. 5. – P. 774–790. DOI: 10.1093/comjnl/bxz019
23. Halder S. A Robotic Cyber-Physical System for Automated Reality Capture and Visualization in Construction Progress Monitoring / S. Halder, K. Afsari, A. Akanmu [Electronic resource]. – Access mode: https://10.48550/arXiv.2402.07034
24. Scalability of Cyber-Physical Systems with Real and Virtual Robots in ROS 2 / [F. J. Mañas-Álvarez, M. Guinaldo, R. Dormido et al.] // Sensors. – 2023. – Vol. 23, Article ID 6073. DOI: 10.3390/s23136073
25. Application of the Computer System Component with Adjustment Elements for Processing Sensor Signals / [H. Ukhina, V. Sytnikov, O. Streltsov et al.] // Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) : 11th IEEE International Conference, Cracow, 22–25 September 2021 : proceedings. – IEEE, 2021. – Vol. 1. – P. 12–17. DOI: 10.1109/IDAACS53288.2021.9660992
26. Quality Factor and Bandwidth [Electronic resource]. – Access mode: https://ecstudiosystems.com/discover/textbooks/basic-electronics/filters/quality-factor-and-bandwidth/
27. Increasing Order of Digital Filters of the Same Type in Robotic Systems / [T. Sytnikov, V. Sytnikov, O. Streltsov et al.] // Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) : 12th IEEE International Conference, Dortmund, 07–09 September 2023 : proceedings. – IEEE, 2023. – Vol. 1. – P. 1200–1204. DOI: 10.1109/IDAACS58523.2023


DOI 10.15588/1607-3274-2024-4-17

УДК 681.326

# ВЕКТОРНО-ЛОГІЧНЕ МОДЕЛЮВАННЯ НЕСПРАВНОСТЕЙ

**Хаханов В. І.** − д-р техн. наук, професор кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**Чумаченко С. В.** – д-р техн. наук, професор, завідувач кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**Литвинова Є. І.** – д-р техн. наук, професор кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**Хаханова Г. В.** – д-р техн. наук, професор кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**Хаханов І. В.** – канд. техн. наук, асистент кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**Рожнова Т. Г.** – канд. техн. наук, доцент кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**Обрізан В. І.** – канд. техн. наук, докторант кафедри автоматизації проектування обчислювальної техніки, Харківський національний університет радіоелектроніки, Україна.

**АНОТАЦІЯ**

**Актуальність.** Технологічні тренди Design&Test комп'ютингу для IT-індустрії та академічної науки завтрашнього дня визначаються такими напрямками: in-memory комп'ютинг, імерсійний комп'ютинг, AI-комп'ютинг, орієнтованими на енергозбереження та скорочення часу обчислень при наданні сервісів. Пропонується механізм моделювання несправностей, як адрес, на розумних структурах даних, які виключають алгоритм моделювання вхідних тестових наборів для отримання тестової карти для логічної функціональності. Запропонований механізм орієнтований на сервісне обслуговування SoC IP-cores під керуванням стандарту IEEE 1500, що може бути сприйнято позитивно інженерами на EDA-ринку.

**Мета.** Мета дослідження – економічні за часом та енерговитратами механізми моделювання несправностей, як адрес, за рахунок використання read-write транзакцій in-memory комп'ютингу для побудови карти тестування будь-якої функціональності на розумних структурах даних.

**Метод.** Розумні структури даних представлені логічним вектором та його похідними у вигляді таблиць істинності та матриць. Карта тестування є матрицею, координати якої визначені комбінаціями всіх логічних несправностей, які перевіряються на двійкових наборах вичерпного тесту. Побудова карти тестування орієнтована на архітектуру in-memory комп'ютингу на основі read-write транзакцій, що робить механізм моделювання економічним до часу моделювання та енерговитрат завдяки відсутності центрального процесора. Логічний вектор як єдиний компонент вхідних даних не вимагає синтезу в технологічно дозволену структуру елементів. Синтез розумних структур даних на основі чотирьох матричних операцій створює карту тестування несправностей, як адрес, для будь-якої логіки.

**Результати.** Вектори дедуктивної матриці ефективно використовуються для моделювання несправностей, як адрес, у цифрових структурах будь-якої конфігурації, включаючи розгалуження, що сходяться, і зворотні зв'язки. Отримана карта тестування використовується для знаходження мінімального тесту перевірки несправностей вхідних змінних. Запропонований механізм моделювання несправностей технологічно легко вписується в архітектуру in-memory комп'ютингу та використовує тільки read-write транзакції. Векторно-логічний механізм можна також використовувати для тестування графових структур, які описуються таблицею істинності або логічним вектором. Адреси таблиці істинності, що використовуються для моделювання несправностей, ефективно застосовуються для безпроцесорної обробки великих даних в архітектурі in-memory комп'ютингу.

**Висновки.** Наукова новизна – пропонується механізм векторно-логічного in-memory комп'ютингу побудови карти тестування, що характеризується побудовою розумних структур даних, які обнулюють алгоритм моделювання несправностей. За простотою та передбачуваністю розмірів структур даних та відсутністю алгоритму моделювання тестових наборів запропонований механізм не має аналогів у design & test індустрії. Практична значимість визначається застосуванням механізму для тестування логічних функціональностей будь-якої складності на вирішення завдань верифікації. Перспективи дослідження – збільшення об'єкта діагностування до схеми, тобто побудова карти тестування схемної логічної структури.

**КЛЮЧОВІ СЛОВА:** Intelligent Computing, In-memory computing, логічний вектор, логічна матриця, карта тестування, структури даних, векторно-логічне моделювання, несправність, таблиця істинності, адреси.

**АБРЕВІАТУРИ**

DAC61 – 61st Design Automation Conference;

EDA – Electronic Design Automation;

SRAM – Static Random Access Memory (статична пам'ять з довільним доступом);

DRAM – Dynamic Random Access Memory (динамічна пам'ять із довільним доступом);

FLASH – Flash Memory (флеш-пам'ять);

RRAM – Resistive Random-Access Memory (резистивна пам'ять із довільним доступом);

PCM – Phase-Change Memory (пам'ять із зміною фазового стану);

MRAM – Magnetoresistive Random-Access Memory (магніторезистивна оперативна пам'ять);


DOI 10.15588/1607-3274-2024-4-18

FinFET-fin field-effect transistor (ребристий польовий транзистор);

AI – Artificial Intelligence (штучний інтелект);

DFX – Design for Excellence (проектування для досконалості);

IFS – Intel Foundry Services (послуги Intel з ливарного виробництва);

SLM – Silicon Lifecycle Management (управління життєвим циклом кремнію);

ATPG – Automatic Test Pattern Generation (автоматична генерація тестових шаблонів);

SoC – System-on-Chip (система на кристалі).

## НОМЕНКЛАТУРА

*L* – логічна матриця;

*Y* – двійковий вектор логічної функціональності;

*F* – карта тестування логічної функціональності (*F*-матриця);

*A* – адреси таблиці істинності (комбінацій логічних несправностей);

*H* – матриці перекодування;

*T* – тестові набори;

*D* – дедуктивна матриця;

$A^1$ – 1-біти;

*n* – кількість змінних;

*w* – довжина регістру (слова).

## ВСТУП

У 2024 році конференція DAC61 [1] зібрала 337 наукових праць з 29 напрямків. Цього року кількість поданих дослідницьких робіт з усіх напрямків та з усього світу зросла на 34%, а також було зафіксовано рекордну кількість поданих 1545 заявок – це після небувалого рекорду за кількістю заявок, встановленого минулого року. Було відзначено, що окрім традиційних тем проектування EDA, IP та вбудованих систем, з'явилися три додаткові теми – штучний інтелект, автономні системи у пам'яті та безпека. Кількість доповідей щодо застосування штучного інтелекту для проектування чіпів, архітектури апаратного та програмного забезпечення за останні кілька років буквально зросла у рази. Тут представлені наукові інновації, що важливо, закінчуються інженерними методиками та додатками.

Для ефективного проектування, тестування, валідації інтегральних схем та компонентів важливо мати просту та розширену мову опису робочого процесу. Сьогодні Python став стандартною мовою програмування для машинного навчання, наукових обчислень та інженерії. Існує також проблема пам'яті. Незважаючи на всі дискусії про закон Мура, одне можна сказати напевно: пам'ять масштабується не так сильно, як логіка. Програми AI, що настільки популярні в наші дні, вимагають все більшого обсягу швидкодіючої та дешевої пам'яті для організації in-memory комп'ютингу та зберігання даних.

До дослідницької програми Research Track було додано програму Engineering Track, призначену для галузевих практиків та технічних менеджерів, яка фокусується на чотирьох ключових областях: інтерфейсне проектування, внутрішнє проектування, інтелектуальна власність, вбудовані системи та програмне забезпечення. Тут інженери галузі діляться останніми інноваціями та ключовими досягненнями. Цього року кількість заявок збільшилась на 32 %, при цьому особлива увага приділялася штучному інтелекту, дизайну та інтелектуальній власності. Декілька десятків доповідей та workshop були присвячені створенню комп'ютерних архітектур у пам'яті. Особливо ця тема звучала при реалізації Artificial Intelligence моделей, економних за енерговитратами та часом. Тут використовуються такі типи пам'яті: SRAM, DRAM, FLASH, RRAM, PCM, MRAM, або FinFET-Nanosheet as their memory technology. Інтеграція AI з in-memory computing надає людству нові можливості економіки великих даних (рис. 1).

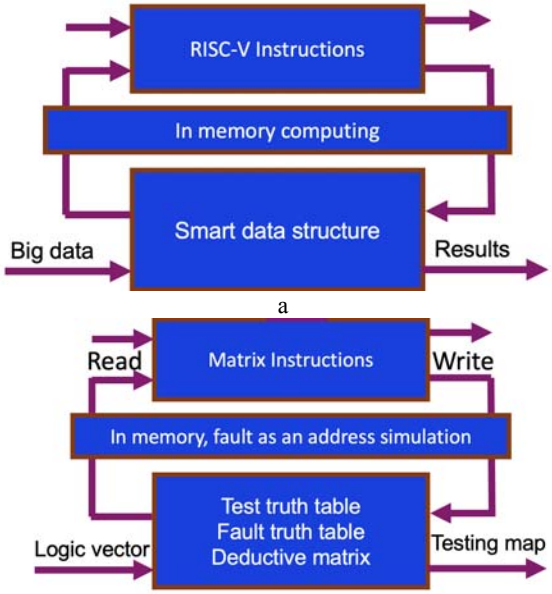



Рисунок 1 – Структура in-memory комп'ютингу обробки даних та моделювання несправностей: а – комп'ютинг у пам'яті; б –моделювання несправності у пам'яті як адрес

На рис. 1 позначено: In-memory computing – комп'ютинг у пам'яті; Smart Data Structure – розумні структури даних; RISC-V Instructions – інструкції обчислювача з набором спрощених/редукованих команд (Reduced instruction set computer); розумні структури даних; Matrix Instructions – матричні інструкції; In-memory, fault as address simulation – моделювання несправності пам'яті як адрес; Test truth table – таблиця істинності тесту; Fault truth table – таблиця істинності несправності; Deductive matrix – дедуктивна матриця; Big Data – великі дані; Results – результати; Read – читання; Write – запис; Logic vector – логічний вектор; Testing map – карта тестування.


DOI 10.15588/1607-3274-2024-4-18

Об'єкт дослідження – in-memory intelligente комп'ютинг, який знижує енергетичні та часові витрати під час обробки великих даних.

Предмет дослідження – in-memory аналіз елементів чи цифрових схем будь-якої розмірності за допомогою read-write транзакцій на логічних векторах.

Будь-які інженерні рішення, орієнтовані на збереження енергії та часу проектування, завжди будуть потрібні на ринку EDA. Виходячи з цього, можна сформулювати *мету дослідження* – економічні за часом та енерговитратами механізми моделювання несправностей, як адрес, за рахунок використання read-write транзакцій in-memory комп'ютингу для побудови карти тестування будь-якої функціональності на розумних структурах даних.

## 1 ПОСТАНОВКА ЗАДАЧІ

Розглядається задача моделювання несправностей як адрес на розумних структурах даних, які виключають алгоритм моделювання вхідних тестових наборів для отримання карти тестування логічної функціональності.

Нехай задано двійковий вектор логічної функціональності *Y*. Слід побудувати карту тестування *F* логічної функціональності *Y*. Практичне завдання полягає у створенні векторно-логічного механізму моделювання несправностей.

Задачі, що підлягають розв'язку: 1) визначення розумних структур даних, що функціонально залежать від логічного вектора; 2) визначення мінімального числа операцій для суперпозиції розумних та явних структур даних, що дозволяють без алгоритму моделювання вхідних тестових наборів визначати карту тестування логічної функціональності; 3) верифікація механізмів моделювання несправностей як адрес, на прикладах логічних функціональностей.

## 2 ОГЛЯД ЛІТЕРАТУРИ

Конференція DAC61 [1] є щорічною помітною подією у світі, на яких обговорюються топ-технології глобального комп'ютингу сьогодення. Це дуже важливо як для коригування програм академічної науки, так і використання AI-пілотів для освоєння курсів студентами [2]. DAC61 23–27 червня 2024 року проходив у Сан-Франциско під прапором «The Chip to Systems», запропонованим компанією Intel. Як зменшити шлях між чіпом та системою для IT-індустрії. Як виготовляти чіпи, орієнтовані на системи. Як швидко заливати системи у виготовлені чіпи та чіплети. Як тестувати та верифікувати системи, що налічують трильйон транзисторів [3]. Крім традиційних тем проектування EDA, IP та вбудованих систем, з'явилися три додаткові важливі теми: штучний інтелект, автономні імерсивні системи та безпека. Кількість доповідей із штучного інтелекту зросла в рази. Вони були присвячені вирішенню питань проектування та верифікації чіпів, апаратного забезпечення та архітектури. DAC61 запропонував трендові чотири чудові доповіді, присвячені штучному інтелекту, створенню обчислювальних систем у пам'яті та іммерсивним обчисленням, а також кілька цікавих виступів SKYtalks та Techtalks, що охоплюють широкий спектр тем, пов'язаних з мікросхемами, системами та додатками. Декілька яскравих прикладів сказаного.

Andrew B. Kahng, Professor UC San Diego: «Останні роки принесли потік рішень щодо штучного інтелекту в EDA. Потенційні переваги AI-EDA включає покращену якість проектування, аналізу та моделювання з меншими часовими, матеріальними та енергетичними витратами [4–11]. Це не залишилося непоміченим як в академічних колах, так і промисловості».

Dr. Gary Patton, Intel, представив концепцію створення «ливарної системи» – чіплети, наповнені програмним забезпеченням. Робиться це шляхом проектування високопродуктивних обчислювачів за епоху штучного інтелекту. Пропозиція зробити екосистему «EDA-IP» має життєво важливе значення для створення 3D-систем на основі штучного інтелекту.

CEO Intel Пат Гелсінгер заявив, що Intel Foundry Services (IFS) відкриває «еру створення систем». Замість того, щоб просто постачати клієнтам пластини, що є традиційною моделлю ливарного виробництва, Intel вже сьогодні пропонує пластини кремнію, корпуси, програмне забезпечення та чіплети. «IFS відкриє еру створення систем, – сказав він, – відзначаючи зміну парадигми, оскільки фокус переміщається від систем на кристалі до «систем-в-корпусі» (SoP) або чіплетам».

Yervant Zorian, Synopsys. Кремній за своєю суттю ненадійний, а пристрої та вузли на його основі найбільш сприйнятливі до відмов. DFX – Design for Excellence на просунутих вузлах потребує нових стратегій. У доповіді розглядалися механізми відмов, методи керування, що забезпечують надійний удосконалений вузол кремнію. Ключові теми включають безпеку, захищеність, надійність та SLM – Silicon Lifecycle Management. Методології та структури управління життєвим циклом кремнію, які необхідні для забезпечення надійності обчислювальних систем.

Sarita V. Adve Professor, University of Illinois. Імерсивні чи просторові обчислення – це загальна концепція низки технологій, які оцифровують діяльність машин, механізмів, людей та об'єктів, а також середовище, в якому вони відбуваються (рис. 2). Створення центру іммерсивних обчислень IMMERSE, який поєднує іммерсивні технології, програми та людський досвід. Показано проект ILLIXR, який базується на наскрізній системі з відкритим вихідним кодом для демократизації


DOI 10.15588/1607-3274-2024-4-18

досліджень іммерсивних систем та їх впровадження до IT-ринку. Імерсивний комп'ютинг вирішує проблеми метричного управління оцифрованими віртуальними, фізичними та соціальними процесами [7–9].

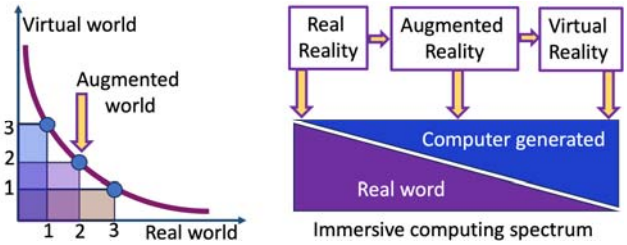



Рисунок 2 – Два подання імерсійного комп'ютингу: Virtual word – віртуальний світ; Augmented world – доповнений світ; Real world – реальний світ; Real Reality – реальна (існуюча) реальність; Augmented Reality – доповнена реальність; Virtual Reality – віртуальна реальність;Computer generated – згенеровано комп'ютером; Immersive computing spectrum – спектр імерсивного комп'ютингу

Alan Lee, Chief Technology Officer, Analog Devices, Штучний інтелект змінює навколишній світ, але основна увага приділяється великим моделям, що працюють на величезних обчислювальних потужностях. Існує гостра потреба в штучному інтелекті в периферійних додатках на основі архітектури in-memory комп'ютингу для зменшення затримок та енергоспоживання [8–11]. Задоволення цієї потреби вимагає нових підходів, що дозволяють задовольнити обмеження майбутніх промислових, автомобільних та споживчих платформ на інтелектуальній периферії. Вартість експлуатації будь-якого комп'ютера – це насамперед вартість електроенергії, спожитої цим пристроєм.

Моделювання несправностей дуже важливе [12–16] для ATPG, верифікації, діагностики, класифікації несправностей. Метрикою ефективності (табл. 1) промислових методів моделювання несправностей є швидкість, пам'ять, складність алгоритмів моделювання функціональних блоків, затримок, елементів, послідовних схем, багатозначне моделювання несправностей, обробка невідомих сигналів *X* та високоімпедансних *Z*.

Враховується складність створення моделей, алгоритмів моделювання елементів та структур. Зручність форми подання моделі справної поведінки та несправностей. Промисловий лідер – спільне моделювання несправностей – це кероване моделювання з використанням good/bad подій. В одному кадрі моделюється підмножина несправних схем, які відрізняються від справної поведінки пристрою. Тут виникають проблеми управління пам'яттю. Практично всі промислові системи моделювання несправностей [14–18] мають непередбачуваний розмір списку несправностей та непередбачуваний розмір структур даних. Усі шість основних методів моделювання використовують процесор із високим рівнем енергоспоживання. Механізм векторного моделювання [19–24] несправностей виграє перед промисловими аналогами по всіх пунктах, крім одного – не враховуються затримки елементів.

## 3 МАТЕРІАЛИ І МЕТОДИ

Суть дослідження, доведеного рівня інженера, полягає у паралельному моделюванні несправностей логіки на вичерпному тесті без алгоритму моделювання вхідних наборів. Інакше кажучи, пропонується адресне моделювання несправностей логіки на інтелектуальних структурах даних без алгоритму моделювання (рис. 3).

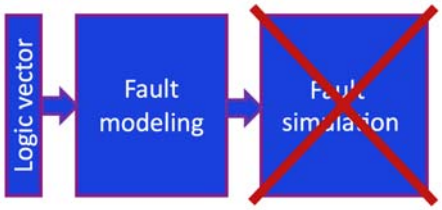



Рисунок 3 – Моделювання несправностей без алгоритму моделювання

Побудова моделі розумних структур даних використовує чотири послідовні процедури синтезу наступних матриць, що формують рішення (рис. 4).

Таблиця 1 – Порівняння промислових технологій моделювання несправностей

| Fault simulation technique | Complexity | Memory | Data structure | Level | Delay | Speed | Fault model | Multy-valued |
|---|---|---|---|---|---|---|---|---|
| Serial fault simulation (1963) | $n \times n^3$ | Predictable | Add fault model, fault list | Gate, system | No problem | Slowest | Any | Easy |
| Parallel fault simulation (1965) | $\frac{1}{w} \times n^3$ | Predictable | Register Memory | Gate | Not capable | Middle | Logic | Dificult |
| Deductive fault simulation (1972) | $n^2$ | Unpredictable | Deductive formulas | Gate | Not capable | Middle | Any | Dificult |
| Concurrent fault simulation (1974) | $\frac{1}{3} \times n^2$ | Unpredictable | Add fault model, fault list | Gate, RTL | Capable | Faster | Logic | Easy |
| PPSFP – Parallel pattern single fault propagation (1985) | $\frac{1}{w} \times n^3$ | Unpredictable | Add fault model, fault list | Gate | Capable | Middle | Logic | Easy |
| Differencial fault simulation (1989) | $\frac{1}{2} \times n^2$ | Unpredictable | Add fault model, fault list | Gate, RTL | Not capable | Middle | Any | Dificult |
| Vector fault simulation (2023) | $\frac{1}{2} \times \frac{1}{3} \times n^2$ | Predictable | No, As true-value simulation | Gate, RTL, System | Not capable | Faster | Logic | Capable |


DOI 10.15588/1607-3274-2024-4-18

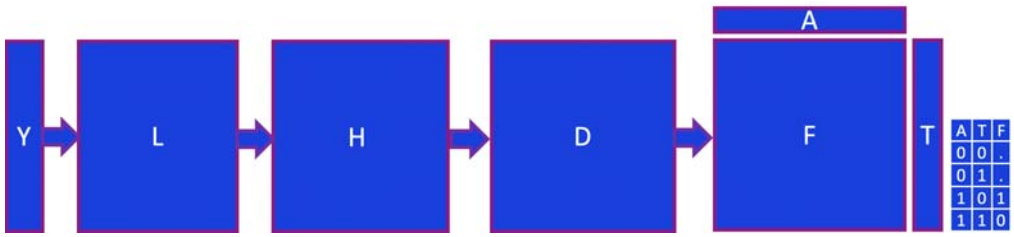



Рисунок 4 – Синтез картки тестування без алгоритму моделювання

1) Синтез логічної *L*-матриці шляхом взяття декартового хог-квадрату на бітах логічного вектора від *n*-змінних за формулою: $L=Y \text{ xor } Y=Y^2_{XOR}$.

2) Побудова матриці перекодування *H* шляхом взяття декартового хог-квадрата на адресах таблиці А істинності від *n* змінних за формулою: $H=A \text{ xor } A=A^2_{XOR}$. Адреси виконують роль тестових наборів *T* та виступають комбінаціями логічних несправностей *A*. Отримана матриця є константою для всіх логічних функцій від *n* змінних.

3) Створення дедуктивної матриці D шляхом переадресації координат логічної *L*-матриці на *H*-матриці перекодування за такою формулою: $D=L_H$.

4) Отримання карти тестування або матриці несправностей, що перевіряються на вичерпному тесті, шляхом виконання координатних операцій на 1-бітах $A^1$ з *A* векторів таблиці істинності в 1-координатах дедуктивної *D*-матриці за формулою: $F=A^1(\neg T)$ або $F=A^1 \text{ xor } T$. Знаки *F* несправностей, що перевіряються, вхідних змінних в координатах карти тестування *F* визначаються інверсією бітів тестових вхідних наборів *T*. Таблиця істинності для формування знаків перевірених несправностей на координатах *F*-матриці наведена в правій частині (див. рис. 4).

## 4 ЕКСПЕРИМЕНТИ

Далі наводиться фрагмент (рис. 5) структури формування знаків несправностей, що перевіряються в алфавіті {0,1, «.»} на координатах карти тестування, відповідних 1-координатам дедуктивної матриці.

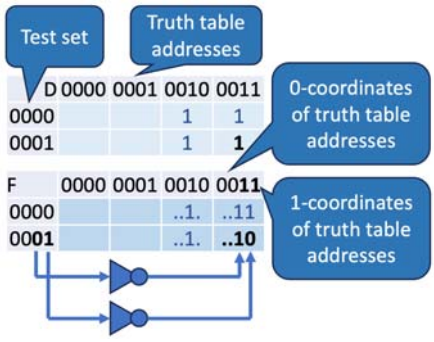



Рисунок 5 – Формування координат карти тестування

Несправності, що перевіряються, формуються тільки по одиничних координатах адрес таблиці істинності. Ці 1-координати визначаються інверсними значеннями бітів двійкових тестових наборів. 0-координати адрес таблиці істинності в карті тестування довизначаються точками, які означають відсутність несправностей, що перевіряються на цих вхідних змінних. Синтез усієї карти тестування функцій від чотирьох змінних буде наведено нижче (рис. 9). Найпростіша реалізація векторно-логічного моделювання несправностей як адрес починається з функціональностей, що мають одну змінну. На основі логічних векторів (00, 01, 10, 11) будуються дедуктивні матриці та карти тестування несправностей (рис. 6). Оскільки логічні вектори 01,10 та 00,11 є взаємно-інверсними, то вони генерують еквівалентні логічні та дедуктивні матриці та карти тестування.

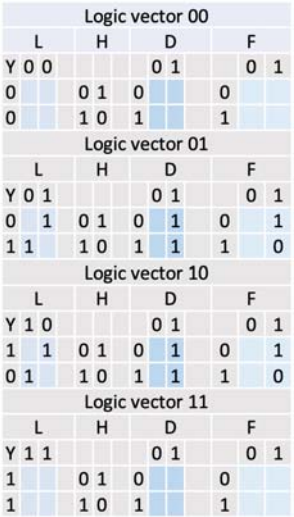



Рисунок 6 – Синтез карт тестування для логіки з однією змінною

Несправності першої та останньої логічної функції не можуть бути перевірені, тому що логічні вектори 00 та 11 не мають змін сигналів. Щоб з'явилися несправності, які можуть бути перевірені, повинні бути логічні вектори, в яких є хоча б один нуль і хоча б одна одиниця.

Наступним пунктом експериментів на механізмі моделювання є побудова карт тестування для кількох логічних функцій двох змінних (рис. 7).

Ці матриці можна використовувати для аналізу несправностей логічних схем з урахуванням стандартних технологічно дозволених елементів. Крім того, їх побудова є навчальною вибіркою для студентів, які бажають освоїти векторно-логічне моделювання несправностей як адрес таблиці істинності логічної функціональності.


DOI 10.15588/1607-3274-2024-4-18

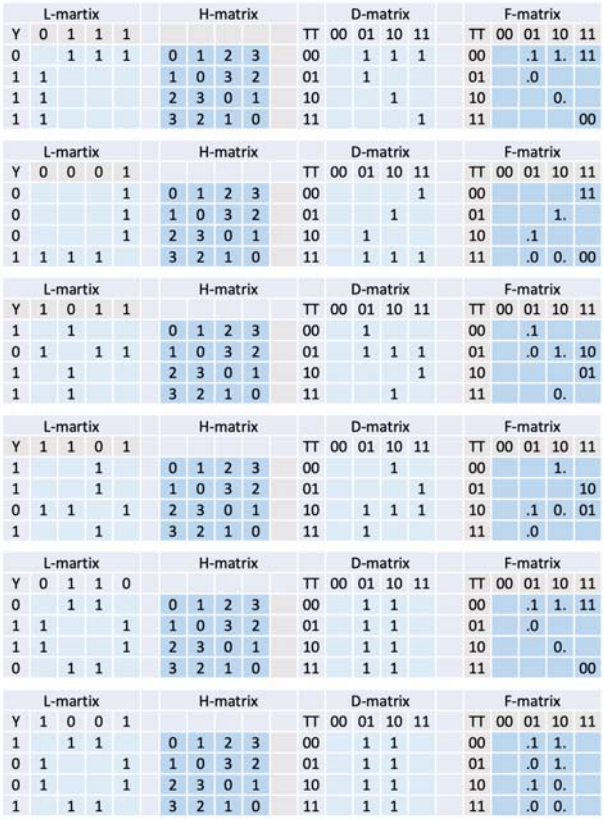


Рисунок 7 – Синтез карт тестування для логіки від двох змінних

Наступним пунктом експериментів на механізмі моделювання є побудова карти тестування для логічних функцій від трьох змінних (рис. 8).

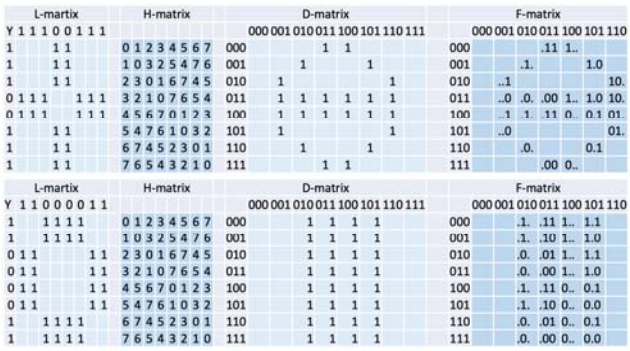


Рисунок 8 – Синтез карт тестування для логіки від трьох змінних

## 5 РЕЗУЛЬТАТИ

Деякі властивості матриць:

1) логічна *L*-матриця завжди симетрична щодо головної діагоналі. Взаємно-інверсні логічні вектори генерують еквівалентні логічні матриці, дедуктивні матриці та карти тестування;

2) кількість одиниць чи активних координат у всіх матрицях є незмінним;

3) вектори *D*-матриці використовуються для моделювання несправностей як адрес при аналізі елементів у цифрових логічних схемах; чим довше *D*-вектор, тим більший ступінь паралелізму обробки несправностей ліній логічної схеми;

4) карта тестування – це повний набір комбінацій вхідних несправностей, що перевіряються на вичерпному тесті. На основі аналізу карти тестування (*F*-матриця) можна легко отримувати мінімальні вхідні тести для перевірки несправностей вхідних змінних розв'язуванням задачі покриття;

5) десяткові коди *H*-матриці – це двійкові коди несправностей, що перевіряються, вхідних змінних або відстані між тестовими вхідними двійковими наборами та адресами таблиці істинності логічної функції.

Заключним пунктом експериментів на механізмі моделювання є побудова карти тестування для логічної функції 00111111111111100 від чотирьох змінних (див. рис. 9).

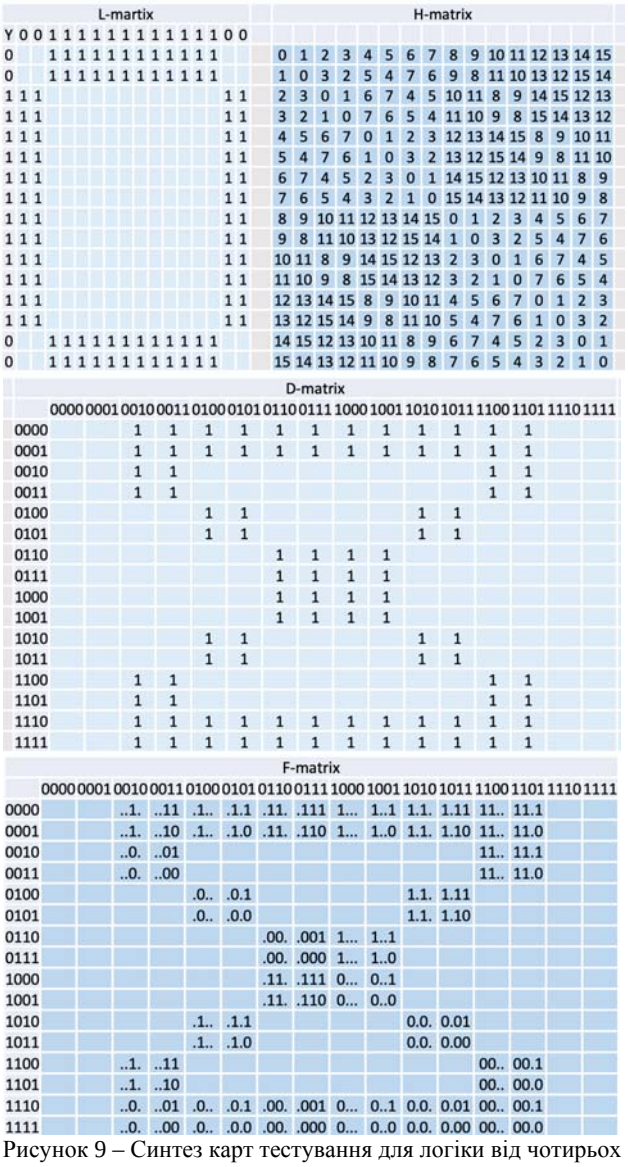


Рисунок 9 – Синтез карт тестування для логіки від чотирьох змінних


DOI 10.15588/1607-3274-2024-4-18

## 6 ОБГОВОРЕННЯ

Перевагою даного векторно-логічного механізму моделювання несправностей як адрес таблиці істинності є новий науковий результат, доведений до рівня інженерного розуміння – як побудувати карту тестування або швидко написати код для її автоматичного синтезу. Основний науковий інноваційний результат – це використання адрес таблиці істинності як тестових вхідних наборів та комбінації логічних несправностей. Судячи з публікацій, цього ніхто досі не робив. Перспектива використання адрес таблиці істинності – це обробка великих даних у пам'яті, де вони зберігаються, за алгоритмом лінійної обчислювальної складності. Обмеження векторно-логічного механізму тестування несправностей пов'язані з візуалізацією матриць і карт тестування від великого (>10) числа логічних змінних.

Ринкова привабливість запропонованого дослідження визначається наступною метрикою, цікавою для EDA-ринку:

1. Економікою обробки великих даних як адрес in-memory комп'ютингу, вільним від процесора [5–7], що дає можливість економити електроенергію (36%) і час обробки великих даних (22%).
2. Суттєвим зниженням обчислювальної складності запропонованих алгоритмів аналізу логічних схем за рахунок експоненційної надмірності розумних структур даних.
3. Інваріантністю запропонованих векторно-логічних структур даних для моделювання несправностей як адрес вентильного, регістрового та системного рівня подання проектів.
4. Передбачуваністю розмірів пам'яті для зберігання розумних структур даних та обчислювальної складності алгоритмів для моделювання несправностей логіки та цифрових схем.
5. Мінімальною кількістю даних про проект для синтезу тестів на основі моделювання логічних несправностей, як адрес.
6. Використання чотирьох векторно-матричних паралельних операцій для синтезу карти тестування логічних функціональностей.
7. Відкриттям тестової константи – універсальної матриці логічних несправностей як відстаней між тестовими наборами та адресами таблиці істинності логічної функції, що у кілька разів зменшує складність алгоритмів моделювання несправностей.
8. Векторно-логічна модель структури дозволяє тестувати несправності переходів будь-якого графа розробленими механізмами синтезу карти тестування. За допомогою кодованого унітарного даного на універсумі примітивів можна побудувати вектор логічної моделі будь-якого соціального процесу, потім його протестувати розробленими механізмами.

## ВИСНОВКИ

Узагальнюючи отримані результати з верифікації векторно-логічного механізму моделювання несправностей, можна зробити такі наукові та практичні висновки: 1) імплементація механізму моделювання в коди мови Python займає 400 рядків; 2) логічного вектора достатньо для побудови карти тестування функціональності, він є суто простою та ефективною інженерною методикою для верифікації SoC IP-cores; 3) звичайно, що питання часу обробки великих за розміром логічних векторів тут не стоїть. Час їхньої обробки – мілісекунди. Проблема полягає лише у візуалізації великих за розміром карт тестування логічних функціональностей, що мають n>10 змінних; 4) вектори дедуктивної матриці ефективно використовуються для моделювання несправностей, як адрес, у цифрових структурах будь-якої конфігурації, включаючи розгалуження, що сходяться, і зворотні зв'язки; 5) отримана карта тестування використовується для знаходження мінімального тесту перевірки несправностей вхідних змінних; 6) даний механізм моделювання несправностей технологічно легко вписується в архітектуру in-memory комп'ютингу і використовує лише read-write транзакції; 7) векторно-логічний механізм можна використовувати для тестування графових структур, які описуються таблицею істинності або логічним вектором; 8) адреси таблиці істинності, що використовуються для моделювання несправностей, ефективно застосовуються для безпроцесорної обробки великих даних в архітектурі in-memory комп'ютингу.

Наукова новизна – пропонується механізм векторно-логічного in-memory комп'ютингу побудови карти тестування, що характеризується створенням розумних структур даних, які обнулюють алгоритм моделювання несправностей.

Практична значимість визначається застосуванням механізму для тестування логічних функціональностей будь-якої складності на вирішення завдань верифікації.

Перспективи дослідження – збільшення об'єкта діагностування до схеми, тобто побудова карти тестування схемної логічної структури.

## ЛІТЕРАТУРА

1. Design Automation Conference «The chips to the systems» [Electronic resource]. Access mode: https://www.dac.com/Portals/0/DAC%2061/Program/DAC_1352254-24_Digital-Onsite-Program-7.pdf?ver=hfG38JtJEimK1HdWaAtD_g%3d%3d
2. News: Professors Rethink. How They Teach Coding Students embrace AI copilots; teachers shift to problem-solving [Text] / [R. D. Caballar, A. Jones, D. Genkina and C. Q. Choi] // IEEE Spectrum. – Vol. 61, № 7. – P. 5–12. DOI: 10.1109/MSPEC.2024.10589683.
3. Liu M. The Path to a 1-Trillion-Transistor GPU: AI's Boom Demands New Chip Technology [Text] / M. Liu and H.-S.


DOI 10.15588/1607-3274-2024-4-18

P. Wong // IEEE Spectrum. – 2024. – Vol. 61, № 7. – P. 22–27. DOI: 10.1109/MSPEC.2024.10589682.
4. MLiM: High-Performance Magnetic Logic in-Memory Scheme with Unipolar Switching SOT-MRAM [Text] / [B. Wu, H. Zhu, K. Chen et al.] // IEEE Transactions on Circuits and Systems I: Regular Papers. – 2023. – Vol. 70, №6. – P. 2412–2424. DOI: 10.1109/TCSI.2023.3254607.
5. Testing for Electromigration in Sub-5nm FinFET Memories [Text] / [M. Mayahinia et al.] // IEEE Design & Test. – 2024. – June. – P. 1. DOI 10.1109/MDAT.2024.3411527.
6. Resistance-Sum Architecture for Voltage-Controlled SOT-MRAM based Computing-in-Memory with Hybrid References [Text] / [C. Xiao et al.] // 2023 IEEE International Magnetic Conference-Short Papers (INTERMAG Short Papers), Sendai, Japan. – 15–19 May 2023: proceedings. – Sendai: IEEE. – P. 1–2. DOI: 10.1109/INTERMAGShortPapers58606.2023.10228265.
7. AI Accelerator Embedded Computational Storage for Large-Scale DNN Models [Text] / [B. Ahn, J. Jang, H. Na et al.] // 2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS), Incheon, Korea, Republic, 13–15 June 2022: proceedings. – Incheon: IEEE, 2022. – P. 483–486. DOI: 10.1109/AICAS54282.2022.9869991.
8. OPC: A Distributed Computing and Memory Computing-Based Efficient Solution of Big Data [Test] / [Z. Yang, C. Zhang, M. Hu, and F. Lin] // 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), Chengdu, China, 19–21 December 2015: proceedings. – Chengdu: IEEE, 2015. – P. 50–53. DOI: 10.1109/SmartCity.2015.46.
9. Reliable ReRAM-based Logic Operations for Computing in Memory [Text] / [M. Moreau et al.] // 2018 IFIP/IEEE International Conference on Very Large-Scale Integration (VLSI-SoC), Verona, Italy, 08–10 October 2018: proceedings. – Verona: IEEE, 2018. – P. 192–195. DOI: 10.1109/VLSI-SoC.2018.8644780.
10. Kang W. Spintronic Memories: From Memory to Computing-in-Memory [Text] / W. Kang, H. Zhang, and W. Zhao // 2019 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH), Qingdao, China, 17–19 July 2019: proceedings. – Qingdao : IEEE, 2019. P. 1–2. DOI: 10.1109/NANOARCH47378.2019.181298.
11. Memory Sizing of a Scalable SRAM In-Memory Computing Tile Based Architecture [Text] / [R. Gauchi et al.] // 2019 IFIP/IEEE 27th International Conference on Very Large-Scale Integration (VLSI-SoC), Cuzco, Peru, 06–09 October 2019: proceedings. – Cuzco: IEEE, 2019. – P. 166–171. DOI: 10.1109/VLSI-SoC.2019.8920373
12. Ulrich E. G. Exclusive Simulation of Activity in Digital Networks [Text] / E. G. Ulrich // Communications of the ACM (CACM). – 1969. – Vol. 13, №2. – P. 102–ll0. DOI: 10.1145/362848.362870
13. Armstrong D. B. A Deductive Method for Simulating Faults in Logic Circuits [Text] / D. B. Armstrong // IEEE Trans. on Computers. – 1972. – Vol. C-21. – P. 464–471. DOI: 10.1109/T-C.1972.223542
14. Ulrich E. G. The Concurrent Simulation of Nearly Identical Digital Networks [Text] / E. G. Ulrich, and T. Baker // 10th Design Automation Workshop, June 25 – 27, 1973: proceedings. – USA: IEEE Press, 1973. – P. 145–150. https://dl.acm.org/doi/epdf/10.5555/800124.804009
15. Donald M. Schuler. Random test generation using concurrent logic simulation [Text] / Donald M. Schuler, Ernst G. Ulrich, Thomas E. Baker, and Susan P. Bryant // 12th Design Automation Conference (DAC '75). – USA: IEEE Press, 1975. – P. 261–267. https://dl.acm.org/doi/abs/10.5555/800261.809076
16. Abramovici M. Digital Systems Testing and Testable Design / M. Abramovici, M.A. Breuer, A.D. Friedman. New York: IEEE Press, 1990. 657 p. DOI:10.1109/9780470544389
17. Chul Young Lee. PROBE: a PPSFP simulator for resistive bridging faults / Chul Young Lee and D.M.H. Walker // 18th IEEE VLSI Test Symposium, Montreal, Quebec, Canada, 30 April 2000 – 04 May 2000: proceedings. – P. 105–110. DOI: 10.1109/VTEST.2000.843833.
18. Riahi, Navabi. A VPI-based combinational IP core module-based mixed level serial fault simulation and test generation methodology [Text] / Riahi, Navabi, and Lombardi // 2003 Test Symposium, Xi'an, China, 16–19 November 2003: proceedings. – IEEE Press^ 2003. – P. 274–277 DOI: 10.1109/ATS.2003.1250822.
19. Vector-logic computing for faults-as-address deductive simulation [Text] / [W. Gharibi, V. Hahanov, S. Chumachenko et al.] // IAES International Journal of Robotics and Automation (IJRA). – 2023. – Vol. 12, № 3. – P. 274–288. DOI: 10.11591/ ijra.v 12i3. P. 274–288.
20. In-Memory Intelligent Computing [Text] / [V.I. Hahanov, V. H. Abdullayev, S. V. Chumachenko, E. I. Lytvynova, I. V. Hahanova] // Radio Electronics, Computer Science, Control. – 2024. – №1. – P. 161–174. DOI: 10.15588/1607-3274-2024-1-15.
21. Vector-Logical Fault Simulation [Text] / [V. Hahanov, S. Chumachenko, E. Litvinova et al.] // Radio Electronics, Computer Science, Control. – 2023. – № 2. – P. 37–51. DOI: 10.15588/1607-3274-2023-2-5.
22. Vector-Driven Logic and Structure for Testing and Deductive Fault Simulation [Text] / [A. Hahanova, V. Hahanov, S. Chumachenko et al.] // Radio Electronics, Comput. Science, Control. – 2021. – №3. – P. 69–85. DOI: 10.15588/1607-3274-2021-3-7.
23. Vector-Deductive Memory-Based Transactions for Fault-as-Address Simulation [Text] / [W. Gharibi, A. Hahanova, V. Hahanov et al.] // Electronic modeling. – 2023. – №45(1). – P. 3–26. DOI: 10.15407/emodel.45.01.003.
24. Vector-Logic Synthesis of Deductive Matrices for Fault Simulation [Text] / [W. Gharibi, A. Hahanova, V. Hahanov et al.] // Electronic modeling. – 2023. – №45(2). – P. 16–33. DOI: 10.15407/emodel.45.02.016.

Стаття надійшла до редакції 17.09.2024.
Після доробки 07.11.2024.


DOI 10.15588/1607-3274-2024-4-18

UDC 681.326

# VECTOR-LOGIC FAULT SIMULATION

**Hahanov V. I. –** Dr. Sc., Professor of the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
**Chumachenko S. V. –** Dr. Sc., Professor, Head of the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
**Lytvynova E. I. –** Dr. Sc., Professor of the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
**Khakhanova H. V.** – Dr. Sc., Professor of the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
**Hahanov I. V.** – PhD, Assistant of the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
**Rozhnova T. G.** – PhD, Associate Professor of the the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
**Obrizan V. I.** – PhD, Post-Doctoral Sudent of the the Design Automation Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

## ABSTRACT

**Context.** The technological trends of Design&Test computing for the IT industry and academic science are determined by the following directions: in-memory computing, immersive computing, AI computing, focused on energy saving and reduction of computing time when providing services. A mechanism for simulating faults as addresses on smart data structures is proposed, which eliminates the algorithm for simulating input test sets to obtain a test map for logic functionality. The proposed mechanism is focused on the service of SoC IP-cores under the control of the IEEE 1500 standard, which can be perceived positively by engineers in the EDA market.

**Objective.** The purpose of the research is time- and energy-saving mechanisms for simulating malfunctions, such as addresses, by using read-write transactions of in-memory computing to build a test map of any functionality on smart data structures.

**Method.** Smart data structures are represented by a logical vector and its derivatives in the form of truth tables and matrices. The test map is a matrix whose coordinates are determined by the combinations of all logical faults that are tested on the binary sets of the comprehensive test. The construction of the test map is focused on the architecture of in-memory computing based on read-write transactions, which makes the simulation mechanism economical in terms of simulation time and energy consumption due to the absence of a central processor. A logical vector as a single component of input data does not require synthesis into a technologically permitted structure of elements. Synthesis of smart data structures based on four matrix operations creates a fault test map like addresses for any logic.

**Results.** Deductive matrix vectors are effectively used to model faults as addresses in digital structures of any configuration, including convergent branches and feedback loops. The resulting test map is used to find the minimum fault-checking test of the input variables. The proposed fault simulation mechanism technologically easily fits into the architecture of in-memory computing and uses only read-write transactions. The vector logic engine can also be used to test graph structures that are described by a truth table or a logical vector. The truth table addresses used for fault simulation are effectively used for processorless processing of large data in the in-memory computing architecture.

**Conclusions.** Scientific novelty – a vector-logic in-memory computing mechanism for building a test map is proposed, characterized by the construction of intelligent data structures that reset the fault modeling algorithm. The proposed mechanism has no analogues in the design & test industry in terms of simplicity and predictability of data structure sizes and the absence of a test set modeling algorithm. The practical significance is determined by the application of the mechanism for testing logical functionalities of any complexity to solve verification tasks. Prospects of the research – increasing the object of diagnosis to the scheme, i.e. building a test map of the scheme logical structure.

**KEYWORDS:** Intelligent Computing, In-Memory Computing, Logic Vector, Logic Matrix, Test Map, Data Structures, Vector-Logic Modeling, Fault, Truth Table, Addresses.

## REFERENCE

1. Design Automation Conference «The chips to the systems» [Electronic resource]. Access mode: https://www.dac.com/Portals/0/DAC%2061/Program/DAC_1352254-24_Digital-Onsite-Program-7.pdf?ver=hfG38JtJEimK1HdWaAtD_g%3d%3d
2. Caballar R. D., Jones A., Genkina D. and Choi C. Q. News: Professors Rethink. How They Teach Coding Students embrace AI copilots; teachers shift to problem-solving [Text], *IEEE Spectrum*, Vol. 61, № 7, pp. 5–12. DOI: 10.1109/MSPEC.2024.10589683.
3. Liu M. and Wong H.-S. P. The Path to a 1-Trillion-Transistor GPU: AI's Boom Demands New Chip Technology [Text], *IEEE Spectrum*, 2024, Vol. 61, № 7, pp. 22–27. DOI: 10.1109/MSPEC.2024.10589682.
4. Wu B., Zhu H., Chen K., Yan C., and Liu W. MLiM: High-Performance Magnetic Logic in-Memory Scheme with Unipolar Switching SOT-MRAM [Text], *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2023, Vol. 70, №6, pp. 2412–2424. DOI: 10.1109/TCSI.2023.3254607.
5. Mayahinia, M. et al. Testing for Electromigration in Sub-5nm FinFET Memories [Text], *IEEE Design & Test*, 2024. June, P. 1. DOI 10.1109/MDAT.2024.3411527.
6. Xiao C. et al. Resistance-Sum Architecture for Voltage-Controlled SOT-MRAM based Computing-in-Memory with Hybrid References [Text], *2023 IEEE International Magnetic Conference-Short Papers (INTERMAG Short Papers), Sendai, Japan, 15–19 May 2023: proceedings.* Sendai, IEEE, pp. 1–2. DOI: 10.1109/INTERMAGShortPapers58606.2023.10228265.


DOI 10.15588/1607-3274-2024-4-18

7. Ahn B., Jang J., Na H., Seo M., Son H. and Song Y. H. AI Accelerator Embedded Computational Storage for Large-Scale DNN Models [Text], *2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS), Incheon, Korea, Republic, 13–15 June 2022: proceedings.* Incheon, IEEE, 2022, pp. 483–486. DOI: 10.1109/AICAS54282.2022.9869991.
8. Yang Z., Zhang C., Hu M., and Lin F. OPC: A Distributed Computing and Memory Computing-Based Efficient Solution of Big Data [Test], *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), Chengdu, China, 19–21 December 2015: proceedings.* Chengdu, IEEE, 2015, pp. 50–53. DOI: 10.1109/SmartCity.2015.46.
9. Moreau M. et al. Reliable ReRAM-based Logic Operations for Computing in Memory [Text], *2018 IFIP/IEEE International Conference on Very Large-Scale Integration (VLSI-SoC), Verona, Italy, 08–10 October 2018: proceedings.* Verona, IEEE, 2018, pp. 192–195. DOI: 10.1109/VLSI-SoC.2018.8644780.
10. Kang W. Zhang H., and Zhao W. Spintronic Memories: From Memory to Computing-in-Memory [Text], *2019 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH), Qingdao, China, 17–19 July 2019: proceedings*. Qingdao, IEEE, 2019, pp. 1–2. DOI: 10.1109/NANOARCH47378.2019.181298.
11. Gauchi R. et al. Memory Sizing of a Scalable SRAM In-Memory Computing Tile Based Architecture [Text], *2019 IFIP/IEEE 27th International Conference on Very Large-Scale Integration (VLSI-SoC), Cuzco, Peru, 06–09 October 2019: proceedings*. Cuzco, IEEE, 2019, pp. 166–171. DOI: 10.1109/VLSI-SoC.2019.8920373
12. Ulrich E. G. Exclusive Simulation of Activity in Digital Networks [Text], *Communications of the ACM (CACM)*, 1969, Vol. 13, №2, pp. 102–ll0. DOI: 10.1145/362848.362870
13. Armstrong D. B. A Deductive Method for Simulating Faults in Logic Circuits [Text], *IEEE Trans. on Computers*, 1972, Vol. C-21. – P. 464–471. DOI: 10.1109/T-C.1972.223542
14. Ulrich E. G. and Baker T. The Concurrent Simulation of Nearly Identical Digital Networks [Text], *10th Design Automation Workshop, June 25–27, 1973: proceedings.* USA, IEEE Press, 1973, pp. 145–150. https://dl.acm.org/doi/epdf/10.5555/800124.804009
15. Donald M. Schuler, Ernst G. Ulrich, Thomas E. Baker, and Susan P. Bryant Random test generation using concurrent logic simulation [Text], *12th Design Automation Conference (DAC '75).* USA, IEEE Press, 1975, pp. 261–267. https://dl.acm.org/doi/abs/10.5555/800261.809076
16. Abramovici M., Breuer M. A., Friedman A. D. Digital Systems Testing and Testable Design. New York, IEEE Press, 1990. 657 p. DOI:10.1109/9780470544389
17. Chul Young Lee and Walker D. M. H. PROBE: a PPSFP simulator for resistive bridging faults, *18th IEEE VLSI Test Symposium, Montreal, Quebec, Canada, 30 April 2000 – 04 May 2000: proceedings*, pp. 105–110. DOI: 10.1109/VTEST.2000.843833.
18. Riahi, Navabi and Lombardi A VPI-based combinational IP core module-based mixed level serial fault simulation and test generation methodology [Text], *2003 Test Symposium, Xi'an, China, 16–19 November 2003: proceedings*, IEEE Press, 2003, pp. 274–277 DOI: 10.1109/ATS.2003.1250822.
19. Gharibi W., Hahanov V., Chumachenko S., Litvinova E., Hahanov I., Hahanova I. Vector-logic computing for faults-as-address deductive simulation [Text], *IAES International Journal of Robotics and Automation (IJRA)*, 2023, Vol. 12, № 3, pp. 274–288. DOI: 10.11591/ ijra.v 12i3. P. 274–288.
20. Hahanov V. I., Abdullayev V. H., Chumachenko S. V., Lytvynova E. I., Hahanova I. V. In-Memory Intelligent Computing [Text], *Radio Electronics, Computer Science, Control*, 2024, №1, pp. 161–174. DOI: 10.15588/1607-3274-2024-1-15.
21. Hahanov V., Chumachenko S., Litvinova E., Hahanova I., Khakhanova A., Shkil A., Rakhlis D., Hahanov I., Shevchenko O. Vector-Logical Fault Simulation [Text], *Radio Electronics, Computer Science, Control*, 2023, №2, pp. 37–51. DOI: 10.15588/1607-3274-2023-2-5.
22. Hahanova A., Hahanov V., Chumachenko S., Litvinova E., Rakhlis D. Vector-Driven Logic and Structure for Testing and Deductive Fault Simulation [Text], *Radio Electronics, Comput. Science, Control*, 2021, №3, pp. 69–85. DOI: 10.15588/1607-3274-2021-3-7.
23. Gharibi W., Hahanova A., Hahanov V., Chumachenko S., Litvinova E., Hahanov I. Vector-Deductive Memory-Based Transactions for Fault-as-Address Simulation [Text], *Electronic modeling*, 2023, №45(1), pp. 3–26. DOI: 10.15407/emodel.45.01.003.
24. Gharibi, W. Hahanova A., Hahanov V., Chumachenko S., Litvinova E., Hahanov I. Vector-Logic Synthesis of Deductive Matrices for Fault Simulation [Text], *Electronic modeling*, 2023, №45(2), pp. 16–33. DOI: 10.15407/emodel.45.02.016.


DOI 10.15588/1607-3274-2024-4-18

# УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

# CONTROL IN TECHNICAL SYSTEMS

UDC 629.072.19

## AN IMPROVED MATHEMATICAL MODEL OF THE METHOD OF FULLY PREPARING THE DETERMINATION OF FIRING UNITS FOR HITTING THE INFORMATION AND CALCULATION COMPONENT OF THE AUTOMATED FIRE CONTROL SYSTEM OF COMBAT VEHICLES OF REACTIVE ARTILLERY

**Majstrenko O. V.** – Dr. Sc., Leading Researcher of Scientific and Methodological Center for the Organization of Scientific and Scientific and Technical Activities National Defence University of Ukraine, colonel, Kyiv, Ukraine.

**Makeev V. I.** – PhD, Associate Professor, Sumy State University, Military Department, colonel, Sumy, Ukraine.

**Prokopenko V. V.** – PhD, Vice Chief of Department of Missile Forces and Artillery Scientific Center of the Army Hetman Petro Sahaidachnyi National Army Academy, Colonel, Lviv, Ukraine.

**Andreiev I. M. –** Sensor Researcher of the Scientific Research Department of the Scientific Center of Missile Forces and Artillery of the National Academy of Ground Forces named after Hetman Petro Sahaidachny, Colonel, Lviv, Ukraine.

**Kamentsev S. Y. –** Leading Researcher of the Scientific Research Department of the Scientific Center of Missile Forces and Artillery of the National Academy of Ground Forces named after Hetman Petro Sahaidachny, Colonel, Lviv, Ukraine.

**Onofriychuk A. Y.** – Junior Researcher of the Research Department of the Scientific Center of the Land Forces of the National Academy of Land Forces Named after Hetman Peter Sahaidachny, Major, Lviv, Ukraine.

**ABSTRACT**

**Context.** As part of the automation of the fire control system of rocket artillery combat vehicles, in relation to the preparation of data for firing and fire control, the information and computing process of this system has been improved, namely, the mathematical model of the method of fully preparing the determination of installations for firing projectiles used in rocket salvo fire systems has been improved. In the system of differential equations of the mathematical model of the information and computing process of the component of the automated fire control system of combat vehicles of jet artillery, weighting functions for air temperature, wind influence for the active and passive sections of the projectile flight trajectory and the section of the opening of combat elements have been introduced, which allows determining the weighting coefficients for them for each projectile type.

**Objective**. To improve the information and calculation component of the automated fire control system of combat vehicles of reactive artillery, by improving the mathematical model of the method of full preparation of the determination of installations for firing on damage. Having proposed a system of differential equations that will take into account the weighting functions of air temperature, wind influence for active and passive sections of the projectile flight path and the section of the opening of combat elements, and will also give the opportunity to determine weighting coefficients for each type of projectile based on them, which in turn will lead to an increase the accuracy of determining firing settings.

**Method**. The proposed analytical method allows: to calculate the weighting coefficients for each type of rocket, characterizing the process of the approach of the rocket flight to the tabular trajectory and to set the initial conditions necessary for solving the differential equations of the mathematical model of the information-computing process of the component of the automated fire control system of combat vehicles of rocket artillery; to increase the accuracy of determining firing positions when performing firing tasks, which makes it possible to quickly respond to a change in the combat situation by means of changes in the software-mathematical process of the automated fire control system; effectively and efficiently ensure the development or clarification of textual and graphic administrative and combat documents based on the results obtained using differential equations of the mathematical model of the information-computational process of the component of the automated fire control system.

**Results**. The improved information and calculation component of the automated fire control system of combat vehicles of jet artillery was tested during the conduct of hostilities. The system of differential equations of the mathematical model of the information-computing process of the component of the automated fire control system of combat vehicles of reactive artillery ensures a timely response to a change in the situation in the information-computational process of the component of the automated fire control system of combat vehicles of reactive artillery during firing and fire control. Provides an opportunity to efficiently and


DOI 10.15588/1607-3274-2024-4-19

quickly ensure the development or clarification of textual and graphic administrative and combat documents based on the information received during the execution of fire missions.

**Conclusions**. The calculations based on the proposed system of differential equations confirm the improvement of the information-calculation component of the automated fire control system of jet artillery combat vehicles and allow timely response to changes in tasks in the information-calculation process during firing and fire control, as well as effectively and quickly ensure the formation of formalized messages and documents based on the information received during the execution of a fire mission by units of reactive artillery. Prospects for further research are the creation of agreed mathematical methods, models, algorithms and programs for the implementation of the goals and tasks of firing and fire control when compiling Firing Tables for prospective or received combat vehicles of reactive artillery from partners.

**KEYWORDS**: automated control system, information and calculation component, mathematical model, system of differential equations, approximation of the functions of the real distribution law, weight functions for air temperature, wind influence.

## ABBREVIATIONS

AFCSRSFS is an automated fire control system of reactive salvo fire systems;

AMS is an automated management system;

ICC is an information and calculation component;

M is a mathematical model;

TAS is a trajectory active section;

TPS is a trajectory passive section;

WFTS is a warhead flight trajectory section;

FP is a full preparation.

## NOMENCLATURE

$a_p$ is a rocket projectile jet acceleration;

$a_r$ is an azimuth of the rocket projectile launch;

$a_x$ is an acceleration of the air drag force of a rocket projectile by TAS;

$B$ is a width of the starting position;

$C_x\left(\frac{V_{r\tau}}{a}\right)$ is an aerodynamic coefficient of the respective projectile;

$C_{X_{ET}}\left(\frac{V_{r\tau}}{a}\right)$ is an aerodynamic coefficient of the reference projectile;

$d$ is a charge caliber;

$E_{X_{ft}^1}$ is a total errors of full preparation by range for the existing method without taking into account the geophysical conditions of firing, combined influence of air temperature, ground atmospheric pressure, wind influence, with entered wind coefficients of direct and cross wind for 9M21HE rockets separately for the indicated sections of the trajectory;

$E_{X_{ft}^2}$ is a total errors of full preparation by range for the existing method without taking into account the geophysical conditions of firing, combined influence of air temperature, ground atmospheric pressure, wind influence, with entered wind coefficients of direct and cross wind for 9M27HE rockets separately for the indicated sections of the trajectory;

$E_{X_{ft}^3}$ is a total errors of full preparation by range for the existing method without taking into account the geophysical conditions of firing, combined influence of air temperature, ground atmospheric pressure, wind influence, with entered wind coefficients of direct and cross wind for 9M55C rockets separately for the indicated sections of the trajectory;

$E'_{X_{ft}^1}$ is a total errors of full preparation by range for the existing method, taking into account the geophysical conditions of firing, the combined effect of air temperature and ground atmospheric pressure, the effect of wind, with the entered wind coefficients of direct and cross wind effect on 9M21HE rockets separately for the indicated sections of the trajectory;

$E'_{X_{ft}^2}$ is a total errors of full preparation by range for the existing method, taking into account the geophysical conditions of firing, the combined effect of air temperature and ground atmospheric pressure, the effect of wind, with the entered wind coefficients of direct and cross wind effect on 9M27HE rockets separately for the indicated sections of the trajectory;

$E'_{X_{ft}^3}$ is a total errors of full preparation by range for the existing method, taking into account the geophysical conditions of firing, the combined effect of air temperature and ground atmospheric pressure, the effect of wind, with the entered wind coefficients of direct and cross wind effect on 9M55C rockets separately for the indicated sections of the trajectory;

$E_{Z_{ft}^1}$ is a total errors of full preparation by direction for the existing method without taking into account the geophysical conditions of firing, the combined influence of air temperature and ground atmospheric pressure, wind influence, with the entered wind coefficients of direct and cross wind influence on 9M21HE rockets separately for the indicated sections of the trajectory;

$E_{Z_{ft}^2}$ is a total errors of full preparation by direction for the existing method without taking into account the geophysical conditions of firing, the combined influence of air temperature and ground atmospheric pressure, wind influence, with the entered wind coefficients of direct and cross wind influence on 9M27HE rockets separately for the indicated sections of the trajectory;

$E_{Z_{ft}^3}$ is a total errors of full preparation by direction for the existing method without taking into account the geophysical conditions of firing, the combined influence of air temperature and ground atmospheric pressure, wind


DOI 10.15588/1607-3274-2024-4-19

influence, with the entered wind coefficients of direct and cross wind influence on 9M55C rockets separately for the indicated sections of the trajectory;

$E'_{Z_{ft}^1}$ is a total errors of full preparation by direction for the existing method, taking into account the geophysical conditions of firing, the combined influence of air temperature and ground atmospheric pressure, wind influence, with the entered wind coefficients of direct and cross wind influence on 9M21HE rockets separately for the indicated sections of the trajectory;

$E'_{Z_{ft}^2}$ is a total errors of full preparation by direction for the existing method, taking into account the geophysical conditions of firing, the combined influence of air temperature and ground atmospheric pressure, wind influence, with the entered wind coefficients of direct and cross wind influence on 9M27HE rockets separately for the indicated sections of the trajectory;

$E'_{Z_{ft}^3}$ is a total errors of full preparation by direction for the existing method, taking into account the geophysical conditions of firing, the combined influence of air temperature and ground atmospheric pressure, wind influence, with the entered wind coefficients of direct and cross wind influence on 9M55C rockets separately for the indicated sections of the trajectory;

$F\left(t, \bar{x}\right)$ is a right part of the system;

$\partial_e$ is an estimated turn on the target from the main direction of fire of the respective firing range, considering meteorological firing conditions;

$\partial_t$ is a turn to the target not considering meteorological conditions;

$F_{58}\left(V_\tau\right)$ is an air resistance law;

$g_0$ is an acceleration of gravity;

$\Delta H$ is a deviation of ground pressure from the table value at the height of the starting position;

$\Delta H_0$ is a ground air pressure at the height of the firing position;

$i_a$ is a rocket projectile shape coefficient by TAS;

$i_n$ is a coefficient of the shape of the rocket projectile by the TPS;

$I_{1N}$ is a table value of a single traction pulse;

$i_{CU}$ is a coefficient of the rocket projectile shape by the WFTS;

$K_1$ is a coefficient that takes into account the effect of the temperature of the reactive charge on a single thrust pulse;

$K_2$ is a coefficient that takes into account the effect of the jet engine's operating time;

$m_0$ is a mass of the rocket projectile;

$q_0$ is a total weight of the projectile;

$q_a$ is a rocket projectile weight by the TAS;

$q_{CU}$ is a rocket projectile weight by TPS;

$q_n$ is a rocket projectile weight by the TPS;

$\Delta q$ is a deviation of the projectile weight from the table value;

$R_E$ is an Earth's radius;

$S$ is a sight of the appropriate firing range taken from the firing tables;

$S_e$ is an estimated sight of the appropriate range of firing rocket projectiles considering meteorological firing conditions;

$\Delta S_{gc}$ is a correction factors of the sight for deviations of geophysical conditions from the table value;

$\Delta S_H$ is a sight correction coefficient for the deviation of the ground atmospheric pressure from the table value at the height of the firing position;

$\Delta S_{H_t}$ is a sight correction coefficient for the deviation of the ground atmospheric pressure from the table value at the height of the target;

$\Delta S_h$ is a sight correction factors for the target exceeding the starting position;

$\Delta S_{T_a}$ is a sight correction factors for ballistic deviation of air temperature from the table value;

$\Delta S_{T_{ah}}$ is a sight correction factors for the combined effect of deviations of air temperature and ground air pressure from the table value;

$\Delta S_{T_{pc}}$ is a sight correction sight correction coefficient for the temperature deviation of powder charges of the main engine;

$\Delta S_{T_{pc_a}}$ is a sight correction coefficient of the sight for the deviation of the temperature of the powder charges of the engine, which corrects the projectile on the trajectory, from the table value;

$\Delta S_{W_{ax}}$ is a correction coefficient of the sight according to the longitudinal component of the ballistic wind within the TAS;

$\Delta S_{W_{nx}}$ is a correction coefficient of the sight according to the longitudinal component of the ballistic wind within the TPS;

$\Delta S_{W_{ex}}$ is a correction coefficient of the sight according to the longitudinal component of the ballistic wind within the WFTS;

$t'$ is a rocket projectile flight time;

$t_a$ is an estimated time of opening of the projectile warhead of the respective firing range, considering the meteorological conditions of firing;

$t_H$ is a jet engine start-up time;

$\Delta t_H$ is a projectile warhead opening time correction factors for deviation of the ground atmospheric pressure from the table value at the height of the starting position;


DOI 10.15588/1607-3274-2024-4-19

$\Delta t_{gc}$ is a projectile warhead opening time correction factors for geophysical conditions deviation from the table value;

$\Delta t_h$ is a projectile warhead opening time correction factors for for exceeding the target over the starting position;

$\Delta t_{T_a}$ is a projectile warhead opening time correction factors for ballistic deviation of air temperature from the table value within the entire trajectory of the projectile;

$\Delta t_{T_{aH}}$ is a projectile warhead opening time correction factors for the combined effect in air temperature and ground air pressure deviations;

$\Delta t_{T_{pc}}$ is a projectile warhead opening time correction factors for deviation of the temperature of main engine powder charges;

$\Delta t_{W_{ax}}$ is a correction coefficient of the opening time of the main part of the projectile according to the longitudinal component of the ballistic wind within the limits TAS;

$\Delta t_{W_{nx}}$ is a correction coefficient of the opening time of the main part of the projectile according to the longitudinal component of the ballistic wind within the limits TPS;

$\Delta t_{W_{ex}}$ is a correction coefficient of the opening time of the main part of the projectile according to the longitudinal component of the ballistic wind within the limits WFTS;

$\Delta T$ is a ballistic air temperature deviation within the full trajectory;

$\Delta T_a$ is a ballistic air temperature deviation;

$\Delta T_a'$ is a ballistic deviation of air temperature from the table value within the entire trajectory of the projectile;

$\Delta T_{pc}$ is a deviation of the temperature of the main engine charges from the table value;

$\Delta T_{pc_a}$ is a deviation of the temperature of the engine charges, which corrects the projectile on the trajectory, from the table value;

$T_{pc}$ is a charge temperature;

$V$ is a rocket projectile flight speed;

$\dot{V}$ is a rocket projectile acceleration;

$V_r$ is a relative speed of rotation of the rocket projectile;

$V_{r\tau}$ is a relative velocity of the projectile, considering the air temperature;

$W_a$ is an average wind speed value by TAS;

$W_{ax}$ is a longitudinal component of the ballistic wind by TAS;

$W_{az}$ is a lateral component of the ballistic wind by TAS;

$W_{CU}$ is an average wind speed by WFTS;

$W_{CU_X}$ is a longitudinal component of the ballistic wind by WFTS;

$W_{CU_z}$ is a lateral component of the ballistic wind by WFTS;

$W_n$ is a wind speed average value by TPS;

$W_{nx}$ is a longitudinal and lateral wind components by TPS;

$W_x$ is a longitudinal component of the ballistic wind;

$W_z$ is a lateral component of the ballistic wind;

$W_{ex}$ is a longitudinal component of the ballistic wind by WFTS;

$X_a$ is a rocket active section projectile flight trajectory length;

$X_n$ is a rocket projectile passive section flight trajectory length;

$X_{ce}$ is a rocket projectile warhead flight trajectory length;

$\dot{x}, \dot{y}, \dot{z}$ is a current value of the velocity changes of the projectile coordinates;

$Y_a$ is a rocket projectile active section flight trajectory final point;

$Y_n$ is a rocket projectile passive section flight trajectory final point;

$Y_s$ is a height of the rocket projectile flight trajectory;

y is a geometrical height;

$\Delta Z_{gc}$ is a correction in the direction of deviation of geophysical conditions from the table value;

$\Delta Z_{W_{az}}$ is a correction factor in the direction of the lateral component of the ballistic wind within the limits TAS;

$\Delta Z_{W_{nz}}$ is a correction factor in the direction of the lateral component of the ballistic wind within the limits TPS;

$\Delta Z_{W_{ez}}$ is a correction factor in the direction of the lateral component of the ballistic wind within the limits WFTS;

$\alpha W_a$ is an average value of the wind directional angle by TAS;

$\alpha W_{CU}$ is an average value of the wind directional angle by WFTS;

$\alpha W_n$ is an average value of the directional wind angle by TPS;

$C_x\left(V_\tau\right)$ is an aerodynamic coefficient of force of frontal air resistance;

$\theta$ is a projectile throwing;

$\dot{\theta}$ is a rate of change of the throwing angle;

$\psi$ is a speed of change of direction of the projectile;


DOI 10.15588/1607-3274-2024-4-19

$\gamma_{aw}$ is a wind coefficient of direct effect of wind on TAS;

$\gamma_{aM}$ is a wind coefficients of the cross effect of the wind on TAS;

$\gamma_{CU_W}$ is a wind coefficient of direct effect of wind on WFTS;

$\gamma_{CU_M}$ is a wind coefficients of the cross effect of the wind on WFTS;

$\mu_y$ is a jet engine of a rocket projectile fuel consumption coefficient;

$\pi(y)$ is an atmospheric pressure with height distribution function;

$\sigma_1$ is an interval on the abscissa axis, on which is defined TAS;

$\sigma_2$ is an interval on the abscissa axis, on which is defined TPS;

$\sigma_3$ is an interval on the abscissa axis, on which is defined WFTS;

$\tau_{aN}$ is a table operating time of the rocket projectile jet engine;

$\Delta\tau$ is a deviation of the virtual air temperature by TAS, TPS, WFTS from table value;

$\tau_y$ is a distribution of virtual temperature with change of height;

$\tau_{ON}$ is a table value of the virtual air temperature on Earth;

$\Psi$ is a rocket projectile angle of penetration;

$\Omega_E$ is an angular velocity of the Earth's rotation;

$\omega_0$ is a rocket projectile weight.

## INTRODUCTION

The analysis of the liberation struggle in Ukraine from the muscovites points to the imperfection of the existing MM of the method of fully preparing the determination of firing units of the ICC of the automated fire control system of combat vehicles of reactive artillery. In real conditions, the calculation using the existing MM of the method of fully preparing the determination of firing units of the ICC of the automated control system will be accompanied by a large number of stages of calculation of weighting factors taking into account ballistic deviations of air temperature and ballistic wind, which will increase the time for the formation of formalized reports and documents on the basis of the received information and will lead to the non-fulfillment of the combat mission due to the irrelevance of the goal. And will also cause a decrease in the accuracy of the projectile falling from the target to 1.6–1.9% in the range and 0–05 – 0–13 divisions of the protractor in the direction [15].

The existing MM of the method of complete preparation of the determination of firing settings of the ICC of the automated control system shows that corrections for the deviation of meteorological, ballistic and geophysical conditions from tabular values are taken into account separately for each section of the projectile flight path, and when calculating corrections for range and direction approximately, correction coefficients are taken into account that characterize the deviation of the projectile in terms of range and direction from the tabular point of fall at all heights of the trajectories (Table 1).

Table 1 – Correction coefficients for calculating corrections in sight and direction

| $R_{km}$ | Correction factors for calculating deviations in | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | sight | | | | | direction | | |
| | $\Delta S_{wax},\ m$ $1\,m/s$ | $\Delta S_{wnx},\ m$ $1\,m/s$ | $\Delta S_{wex},\ m$ $1\,m/s$ | $\Delta S_H,\ m$ $1\ mm.\ m.\ c$ | $\Delta S_{T_o},\ m$ $10^0\ C$ | $\Delta Z_{wax},\ m$ $1\,m/s$ | $\Delta Z_{wnx},\ m$ $1\,m/s$ | $\Delta Z_{wex},\ m$ $1\,m/s$ |
| 30 | 76 | 22 | 70 | 49 | 60 | 15 | 11 | 56 |
| 40 | 46 | 16 | 41 | 45 | 34 | 22 | 26 | 42 |
| 50 | 49 | 48 | 35 | 62 | 14 | 30 | 36 | 37 |
| 60 | 59 | 56 | 42 | 78 | 16 | 39 | 42 | 40 |
| 70 | 68 | 55 | 52 | 90 | 53 | 47 | 43 | 45 |

Analyzing the data presented in Table 1, it can be concluded that the effect of the longitudinal wind on the movement of the projectile in the active section of the trajectory (TAS), the passive section of the trajectory (TPS) and the section of the flight trajectory of combat elements (WFTS) is approximately the same, with the exception of distances from 30 up to 40 km, where the influence of the bucket on the PDT is 2–2.5 times less compared to its influence within the ADT and on the DTPBE, and the effect of the side wind within the ADT and PDT is identical at all firing ranges, except for the DTPBE in 2–3 times higher than the impact on ADT and PDT at distances of 30–40 km.

## 1 PROBLEM STATEMENT

The main shortcomings of the MM of the method for the complete preparation of determination of settings for firing the ICC of the automated control system during the calculation of the values of the settings for defeating individual and lightly armored targets are: the accuracy of determining the median deviations in range and direction; failure to take into account the weighting factors for accounting for ballistic deviations of air temperature and ballistic wind on the active and passive sections of the trajectory, as well as on the flight section of combat elements. And also the imperfection of: mathematical (numerical) methods of solving problems in the ICC of the automated control system regarding the approximation


DOI 10.15588/1607-3274-2024-4-19

of the functions of the real law of distribution [2, 7, 12]; differential equations of the MM of the method of full preparation for determining the installations for shooting, the information-calculating component of the automated control system for determining the meteorological, ballistic, topographic and geophysical conditions of shooting [1, 6]. An important direction of the accuracy of the calculation of the correction values for the deviation of the meteorological conditions of shooting from the tabular values, using a MM of the method of full preparation of the determination of the settings for shooting information-calculation component of the automated system management, there is an approach based on the approximation of the functions of the real distribution law of the deviations of meteorological factors from tabular values, which can be described by the corresponding distribution ratios expressed by the algebra of intervals [18, 20].

The existing approaches to the approximation of the functions of the real law of the distribution of deviations of meteorological factors from tabular values in the system of differential equations of the MM of the information-computing process of the component of the automated fire control system of combat vehicles of reactive artillery have a number of significant shortcomings.

The main ones are: replacing real distribution laws with linear ones, which leads to significant errors in the calculation of meteorological corrections and is approximate; increased efficiency of calculating the values of corrections for the deviation of meteorological shooting conditions from tabular values, which significantly affects the calculation time, and, accordingly, the efficiency and speed of forming formalized messages and documents based on the received information; separate accounting of meteorological conditions on the sections of the trajectory, which significantly complicates and reduces the accuracy of calculations of weighting factors to take into account ballistic deviations of air temperature and ballistic wind [6, 9, 10].

The existing MM of the method for the complete preparation of the determination of firing positions of the ICC of the automated fire control system of jet artillery combat vehicles, namely, the system of differential equations does not take into account the weighting factors on the active, passive sections of the trajectory and the section of the flight of combat elements, which leads to errors during calculation of firing settings and, as a result, affects the deviation of the projectile from the target in terms of range and direction by up to 5–7%. And this significantly reduces the efficiency and speed of forming formalized messages and documents on the basis of the received information and, accordingly, defeating the target (Table 2).

Table 2 – Deviation of projectiles by range and direction at the values of the determined meteorological conditions

| $R_{km}$ | in sight, m | | | | | in direction, m | | |
|---|---|---|---|---|---|---|---|---|
| | $W_{a_x}=15\ m/s$ | $W_{n_x}=20\ m/s$ | $W_{e_x}=15\ m/s$ | $\Delta H_0=20\ mm.m..c$ | $T_a=20^0\,C$ | $W_{a_z}=15\ m/s$ | $W_{n_z}=20\ m/s$ | $W_{e_z}=15\ m/s$ |
| 30 | 1140 | 440 | 1050 | 980 | 1200 | 225 | 220 | 840 |
| 40 | 690 | 320 | 615 | 900 | 680 | 330 | 520 | 630 |
| 50 | 735 | 960 | 525 | 1240 | 280 | 450 | 720 | 555 |
| 60 | 885 | 1120 | 630 | 1560 | 320 | 585 | 840 | 500 |
| 70 | 1020 | 1100 | 780 | 1800 | 1060 | 705 | 860 | 675 |

Let’s consider one of the components of the information and calculation process of the AMS. The component is related to the determination of the settings for shooting, the calculation of the values of corrections for the deviation of meteorological conditions from the tabular conditions by evaluating the proposed method of their determination.

Calculation of settings for firing is carried out using mathematical dependencies, taking into account the tabular deviation values: ground atmospheric pressure at the height of the firing position $(\Delta H_0)$, ballistic air temperature deviation within the full trajectory $(\Delta T)$, ballistic air temperature deviation $(\Delta T_a)$, longitudinal and lateral components of the ballistic wind $(W_x, W_z)$.

In the MM of the method of complete preparation of the determination of firing units for hitting the ICC of the automated fire control system of combat vehicles of jet artillery, during the calculation of the values of corrections for the deviation of meteorological conditions from the tabular ones, the flight path of the jet projectile is divided into three sections: trajectory active section; trajectory passive section; warhead flight trajectory section (Fig. 1).

Therefore, it is necessary to calculate the values of the calculated corrections, by range and direction, for the deviation of meteorological conditions from tabular values, taking into account geophysical conditions: the deviation of the temperature of charges on the trajectory, the combined effect of air temperature and ground atmospheric pressure, as well as the effect of wind to be determined by dependences (1, 2, 3), for each section of the trajectory [2, 12, 15].

## 2 REVIEW OF THE LITERATURE

At present, during the experiments of projectile flight research using a MM of the method of fully preparing the determination of firing units of the ICC of the automated fire control system of jet artillery combat vehicles, it is very relevant to study the consideration of corrections for the deviation of meteorological, ballistic and geophysical


DOI 10.15588/1607-3274-2024-4-19

firing conditions from the tabular values separately for each section of the flight path of this projectile. In general, during calculations, with the help of a MM of the method of full preparation of the determination of firing units of the ICC of the automated fire control system of combat vehicles of jet artillery, the values of corrections for range and direction, taking into account the correction coefficients characterizing the deviation of the projectile in terms of range and direction from the table the drop points at all heights of the projectile flight path are approximate, which affects the accuracy of hitting targets and is 1.6–1.9% in range and 0–05 – 0–13 protractor divisions in direction [3, 11, 19].

$$S_e = S + \Delta S_{W_{ax}} W_{ax} + \Delta S_{W_{nx}} W_{nx} + \Delta S_{W_{ex}} W_{ex} + \Delta S_H \Delta H + \Delta S_{H_\tau} \Delta H^2 + \Delta S_{T_a} + \\ + \Delta T_a + \Delta S_{T_{pc}} T_{pc} + \Delta S_{T_{pc_e}} \Delta T_{pc_e} + \Delta S_{gc} + \Delta S_{T_{ah}} + \Delta T_a^2 + \Delta S_h + \Delta S_{T_H} \Delta T' \Delta H \quad , \tag{1}$$

$$t_a = t + \Delta t_{W_{ax}} W_{ax} + \Delta t_{W_{nx}} W_{nx} + \Delta t_{W_{ex}} W_{ex} + \Delta t_H \Delta H + \Delta S_{T_a} \Delta T_a + \Delta t_{T_{pc}} T_{pc} + \Delta t_{gc} + \Delta t_h + \Delta t_{T_{aH}} \Delta T_a \Delta H \ , \tag{2}$$

$$\partial_e = \partial_t + \Delta Z_{W_{az}} W_{az} + \Delta Z_{W_{nz}} W_{nz} + \Delta Z_{W_{ez}} W_{ez} + \Delta Z_{gc} \tag{3}$$

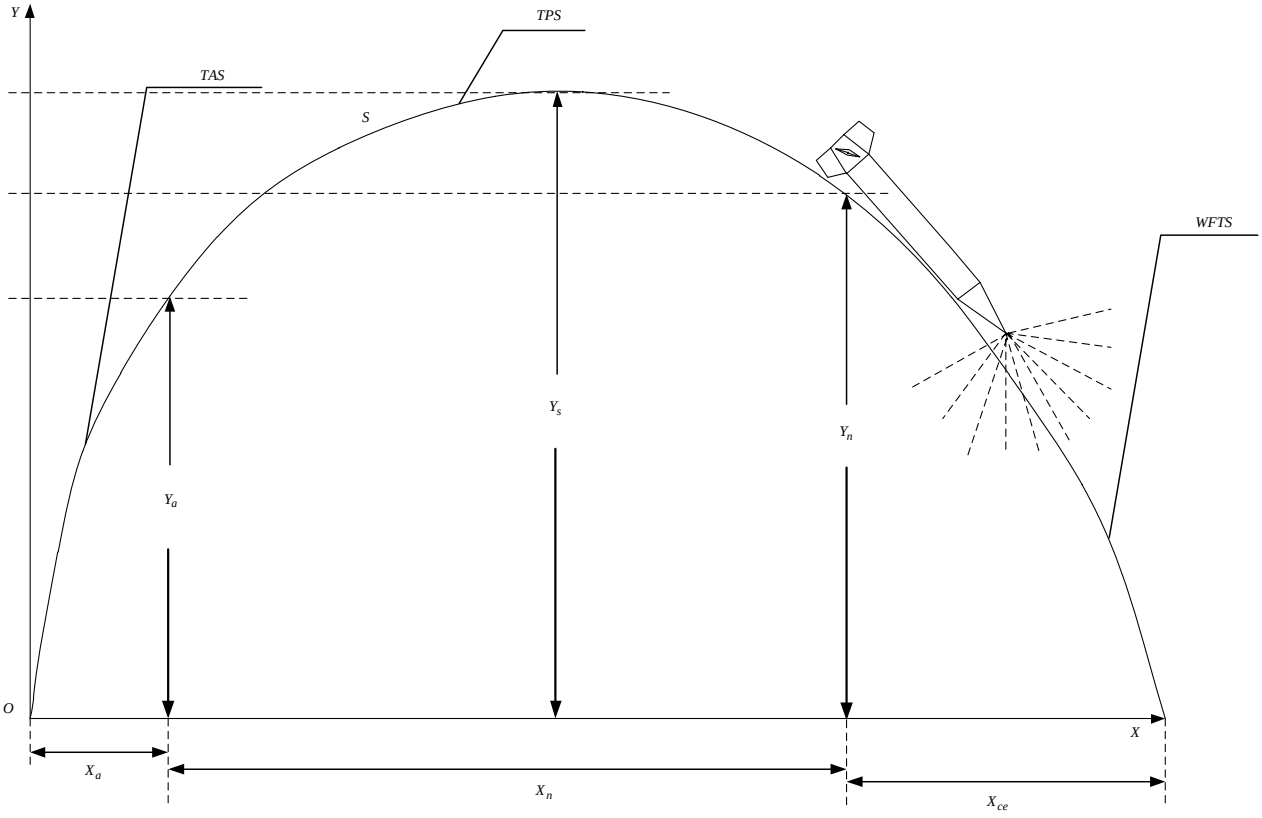



Figure 1 – Rocket projectile flight trajectory

In modern ballistics literature, there are no MM of the method of fully preparing the determination of firing settings of the ICC of the automated fire control system of combat vehicles of reactive artillery [1–3, 16–18] which would make it possible to calculate firing installations taking into account corrections for the deviation of meteorological firing conditions from tabular conditions separately for each section of the missile flight path.

The paper [8] describes a new multidisciplinary computational study conducted to simulate the flight trajectories and free-flight aerodynamics of both a finned projectile at supersonic speeds and a rotating projectile at subsonic speeds with and without aerodynamic flow control. The method of effective formation of a complete aerodynamic description for dynamic simulation of projectile flight is described in [16].

The textbook [13] contains basic information about the motion of an aerodynamic body, and gives the equations of motion of aircraft for the active and passive sections of the trajectory. In works [16, 17], a model of a mobile solid-fueled aircraft is proposed, taking into account the rotation and curvature of the Earth, the influence of wind and other aerodynamic parameters; the study of the impact of ramjet operation parameters on the


DOI 10.15588/1607-3274-2024-4-19

range and accuracy of rockets was carried out in [14]. A realistic nonlinear model of flight dynamics was developed to perform simulations to confirm the accuracy of the presented algorithms [4, 5]. The analysis of projectile flight dynamics with full six degrees of freedom is considered in [1, 17]. The article [14] presents a new approach to controlling the external ballistic properties of spin-stabilized bullets by optimizing their internal mass distribution. Currently, MM for the study of external ballistics are cumbersome. This work aims to define and expand the theoretical procedure for determining the complex of indicators that accompany the movement of the projectile, as well as the MM of the method of fully preparing the determination of firing units for the ICC of the automated fire control system of combat vehicles of reactive artillery, which is presented in the work and can be used for calculating the values of corrections for the deviation of the meteorological conditions of firing from the tabular conditions in the firing units separately for each section of the missile flight path.

## 3 MATERIALS AND METHODS

To determine the correction factors for TAS, TPS and WFTS considering the geophysical conditions of firing, the combined effect of air temperature and ground pressure of the atmosphere, as well as the wind effect, we will use the known system of differential equations of the MM of the method of complete preparation of the determination of firing units of the ICC of the automated fire control system of combat vehicles of reactive artillery [5], but when calculating the relative velocity of the projectile and the acceleration of the air drag force, we will introduce wind coefficients of range and cross wind effects on projectiles separately for the specified sections of the trajectory (4).

$$\begin{cases} \dot{x} = V\cos\theta\cos\psi \Big/ \left(1-\dfrac{2y}{R_E}\right); \\ \dot{y} = V\sin\theta; \\ \dot{z} = V\cos\theta\sin\psi; \\ \dot{V} = a_p - a_x\cos\gamma - g_0\sin\theta\left(1-\dfrac{2y}{R_E}\right); \\ \dot{\theta} = \dfrac{\cos\theta g_0\left(1-2y/R_E\right)}{V} - \dfrac{a_x\cos\gamma W_x\sin\theta}{VV_r} + \dfrac{V\cos\theta}{R_E+y} - \\ -\Omega_E\cos B\sin\left(a_{r-\psi}\right); \\ \dot{\psi} = -\dfrac{a_x\cos\gamma W_z}{\cos\theta VV_r} + \Omega_E\left(\sin B_l - \cos B_l\cos\left(a_{r-\psi}\right)tg\theta\right); \\ \dot{\pi}(y) = -\dfrac{\pi(y)\dot{y}}{R\left[\tau_y+\Delta\tau\right]} \end{cases}, \quad (4)$$

where, $a_p = \dfrac{\omega_0\left(I_{1N}+K_1\Delta T_{pc}\right)}{m_0\left[\tau_{aN}-K_2\Delta T_{pc}\right]\left(1-\mu_y\right)}$,

$\Delta T_{pc} = T_{pc} - 15$, $m_0 = q_0 \Big/ g_0$,

$\mu_y = \dfrac{\omega_0\left(t'-t_H\right)}{g_0 m_0\left(\tau_{aN}+K_2\Delta T_{pc}\right)}$,

$a_x = 0.474\dfrac{i_a d^2}{q_0}10^3\dfrac{\tau_{ON}}{\tau_y+\Delta\tau}\dfrac{\pi(y)F_{58}\left(V_{r\tau}\right)}{1-\mu_y}$,

$i_a = C_x\left(\dfrac{V_{r\tau}}{a}\right) \Big/ C_{X_{ET}}\left(\dfrac{V_{r\tau}}{a}\right)$, $V_{r\tau} = V_r\sqrt{\dfrac{\tau_{ON}}{\tau_y+\Delta\tau}}$

$$V_r = V\sqrt{1-\frac{2W_{ax}\left(\cos\theta\cos\psi+\gamma_{aw}W_{ax}\sin\psi-\gamma_{aM}W_{az}\sin\theta\cos\psi-\gamma_{aw}\gamma_{aM}W_{ax}W_{az}tg\theta\cos\psi\right)}{V} - \frac{2W_{ax}\left(\cos\theta\sin\psi-\gamma_{aw}W_{ax}\cos\psi-\gamma_{aM}W_{az}\sin\theta\sin\psi+\gamma_{aw}\gamma_{aM}W_{ax}W_{az}tg\theta\cos\psi\right)}{V} + \frac{W_a^2}{V^2}}$$

at $\theta\to\theta_0+\delta\theta,\ \psi=\psi_0+\Delta\psi$, $W_a^2 = W_{ax}^2 + W_{az}^2$, $W_{ax} = W_a\cos\alpha W_a$, $W_{az} = W_a\sin\alpha W_a$;

$$V_r = V\sqrt{1-\frac{2W_{ax}\left(\cos\theta\cos\psi+\gamma_{aw}W_{ax}\sin\psi-\gamma_{aM}W_{az}\sin\theta\cos\psi+\gamma_{aw}\gamma_{aM}W_{ax}W_{az}tg\theta\sin\psi\right)}{V} - \frac{2W_{ax}\left(\cos\theta\sin\psi+\gamma_{aw}W_{ax}\cos\psi-\gamma_{aM}W_{az}\sin\theta\sin\psi-\gamma_{aw}\gamma_{aM}W_{ax}W_{az}tg\theta\cos\psi\right)}{V} + \frac{W_a^2}{V^2}}$$

if $\theta\to\theta_0+\delta\theta,\ \psi=\psi_0-\Delta\psi$;


DOI 10.15588/1607-3274-2024-4-19

$$V_r = V\sqrt{1-\frac{2W_{ax}\left(\cos\theta\cos\psi-\gamma_{aw}W_{ax}\sin\psi+\gamma_{aM}W_{az}\sin\theta\cos\psi-\gamma_{aw}\gamma_{aM}W_{ax}W_{az}tg\theta\sin\psi\right)}{V}-\frac{2W_{ax}\left(\cos\theta\sin\psi+\gamma_{aw}W_{ax}\cos\psi+\gamma_{aM}W_{az}\sin\theta\sin\psi+\gamma_{aw}\gamma_{aM}W_{ax}W_{az}tg\theta\cos\psi\right)}{V}+\frac{W_a^2}{V^2}}$$

if $\theta\to\theta_0-\delta\theta,\ \psi=\psi_0-\Delta\psi$;

$$V_r = V\sqrt{1-\frac{2\left(W_{nx}\cos\theta\cos\psi+W_{nx}\cos\theta\sin\psi\right)}{V}+\frac{W_n^2}{V^2}},\quad W_n^2=W_{nx}^2+W_{nz}^2,\quad W_{nx}=W_n\cos\alpha W_n,\ W_{nz}=W_n\sin\alpha W_n;$$

$$V_r = V\sqrt{1-\frac{2W_{CU_x}\left(\cos\theta\cos\psi+\gamma_{CU_w}W_{CU_x}\sin\psi-\gamma_{CU_M}W_{CU_z}\sin\theta\cos\psi-\gamma_{CU_w}\gamma_{CU_M}W_{CU_x}W_{CU_z}tg\theta\sin\psi\right)}{V}-\frac{2W_{CU_x}\left(\cos\theta\sin\psi-\gamma_{CU_w}W_{CU_x}\cos\psi-\gamma_{CU_M}W_{CU_z}\sin\theta\sin\psi+\gamma_{CU_w}\gamma_{CU_M}W_{CU_x}W_{CU_z}tg\theta\cos\psi\right)}{V}+\frac{W_{CU}^2}{V^2}}$$

where $W_{CU}^2=W_{CU_x}^2+W_{CU_z}^2$, $W_{CU_x}=W_{CU}\cos\alpha W_{CU}$, $W_{CU_z}=W_{CU}\sin\alpha W_{CU}$,

$$\tau_y=\begin{cases}289.0-0.006328Y\ if\ 0\le Y\le 9324,\\ 230-0.006328\left(Y-9324\right)+0.000001172\times\\ \times\left(Y-9324\right)^2\ if\ 9324\le Y\le 12000,\\ 221.5\ if\ Y\rangle 12000.\end{cases}$$

Accordingly, in the system of differential equations (4), the acceleration of the air drag force by the TPS will be determined by the formula

$$a_x = 0.474\frac{i_n d^2}{q_n}10^3\frac{\tau_{ON}}{\tau_y+\Delta\tau}F_{58}\left(V_{r\tau}\right),\qquad(5)$$

where $q_n=q_a-\omega_0$, and the acceleration of the air drag force by the WFTS will be determined by equation

$$a_x = 0.474\frac{i_{CU} d^2}{q_{CU}}10^3\frac{\tau_{ON}}{\tau_y+\Delta\tau}F_{58}\left(V_{r\tau}\right).\qquad(6)$$

## 4 EXPERIMENTS

We will give the system (4) in a kind

$$\frac{d\bar{x}}{dt}=F\left(t,\bar{x}\right).\qquad(7)$$

We consider that system of nonlinear differential equalizations of external ballistics of kind (5) relatively variables $\bar{x}\in\Re^n$, $t\in I=[0,\infty)$, thus $\bar{x}\left(x,y,z,V,\theta,\psi,\pi(y)\right)$ is vector-line of variables $(n=6)$, fulfills conditions of existence, uniqueness and continuous dependence on the initial conditions of decisions in a region [20] $\left(\bar{tx}\right)\subset G_H=IB_h$; $B_h=\left\{\bar{x}\in\Re^n:\|\bar{x}\|\langle H\right\}$.

Then taking into consideration essence of the use of procedure, which consists in that in the process of calculation of relative speed of reactive to the shell and we will enter acceleration of force of head-resistance of air winds coefficients of the direct and cross influencing of force wind on jet-projectiles separately for the indicated areas of trajectory as a result providing of exactness of implementation of this procedure is possible by bringing in of principle of Borel-Lebeg [20], in obedience to which it is possible with a confidence to consider that in the case of reserved to the segment on an axis OX it is always possible to select a complete subsystem which fully will cover this segment; hereupon complete trajectory of flight to the shell, certain on some complete interval, will be covered by the complete set of intervals, each of which will answer the separate areas of complete trajectory of motion of shell.

Consequently, reserved interval $\lfloor 0; X_a+X_p+X_{fe}\rfloor$, which a curve, that is associated with flight to shell, is certain on, always will be covered by some system of segments $\sum=\{\sigma\}$ of the opened intervals which a complete subsystem is selected from

$$\sum{}^*=\{\sigma_1,\sigma_2,\sigma_3\},\qquad(7)$$

where $\sigma_1=T_{TAS}=[0;X_a]$, $\sigma_2=T_{TPS}=\left[X_a;X_a+X_p\right]$, $\sigma_3=T_{WFTS}=\left[X_a+X_p;X_a+X_p+X_{fe}\right]$ – segments of line on abscises axis, on which certainly active, passive and opening of battle elements accordingly.


DOI 10.15588/1607-3274-2024-4-19

## 5 RESULTS

The results of calculations of the values of the total average errors by range and direction using a system of differential equations (4) of the MM of the method of complete preparation of the determination of firing units of the information-calculation component of the automated fire control system of combat vehicles of jet artillery, considering the geophysical conditions of firing, the combined effect of air temperature and ground pressure, as well as the effect of wind, with the introduction of wind coefficients of direct and cross wind effects on rockets separately for the specified trajectory sections for 122-mm combat vehicle BM-21 –9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile and 300-mm 9A52 9M55C fragmentation projectile are given in Table 3.

From the data presented in the table, it can be concluded that the calculations of the total average errors by range and direction when determining firing settings by the FP method, using the proposed system of differential equations for 122-mm combat vehicle FM-21 – 9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile and 300-mm 9A52 9M55C fragmentation projectile are reduced by half during firing.by range and direction when determining firing.
Calculations of the total FP errors according to the system (4) in the range for 122-mm combat vehicle FM-21 – 9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile and 300-mm 9A52 9M55C fragmentation projectile are given on the Fig. 2.

The following conclusions can be drawn from the given data: direction when determining the firing settings by FP method for accuracy depends on the firing range and the multiple launch rocket system; when firing at medium and long ranges, the value of the total average range errors is reduced by 1.5 times, taking into account the geophysical conditions of firing, the combined effect of air temperature and ground pressure, as well as the effect of wind, with the wind coefficients of direct and cross wind effects on rockets separately introduced for the specified trajectory sections; taking into account the geophysical conditions of firing, the combined effect of air temperature and ground pressure, as well as the effect of wind, with the introduced wind coefficients of direct and cross wind effects on rocket projectiles separately for the specified trajectory sections, the total average directional errors at the minimum firing ranges almost do not change and make up a difference of 0–01 for 122-mm combat vehicle BM-21 – 9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile, and for 300-mm 9A52 projectile 9M55C is halved; taking into account the geophysical conditions of firing, the combined effect of air temperature and ground pressure, as well as the effect of wind, with the introduced wind coefficients of direct and cross wind effects on rockets separately for the specified trajectory sections, the total average errors in the direction at maximum firing ranges for 122-mm combat vehicle BM-21 – 9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile, and for 300-mm 9A52 projectile 9M55C is halved.

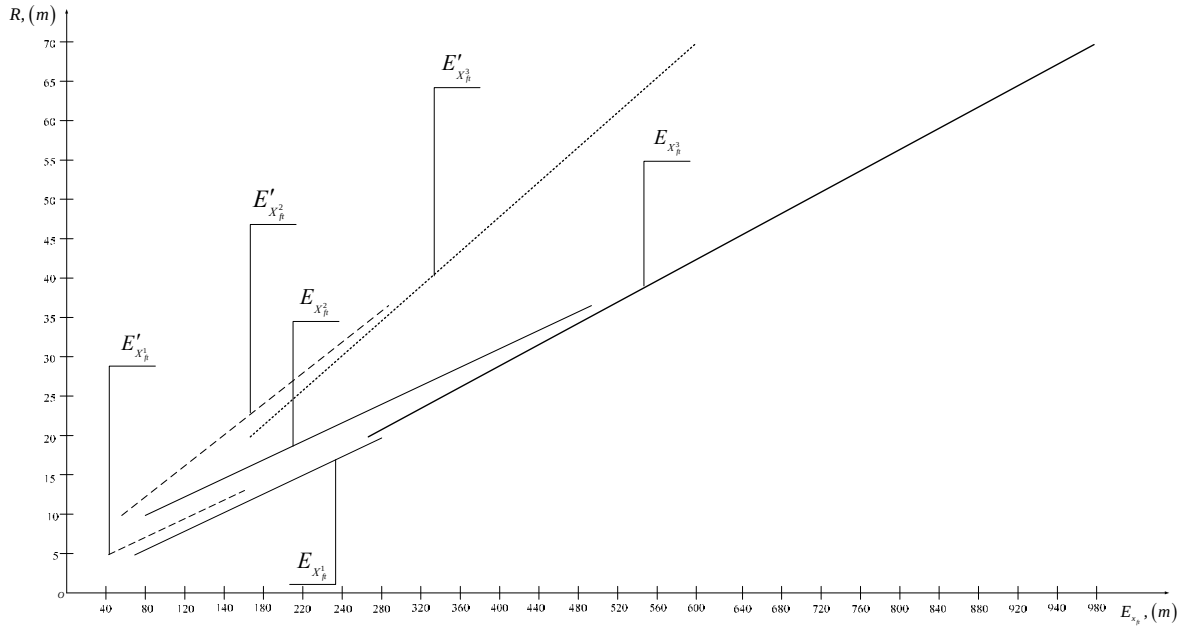
## 6 DISCUSSION

Determination of correction coefficients on the active and passive sections of the trajectory and on the section of the flight trajectory of combat elements, taking into account the geophysical conditions of firing, the combined effect of air temperature and ground atmospheric pressure, as well as the effect of wind, is carried out using a system of differential equations of the mathematical model of the method of complete preparation of the determination of firing installations of the information and calculation component of the automated fire control system of combat vehicles of reactive artillery [2, 13, 17], or by the Dmitrievsky method [21].

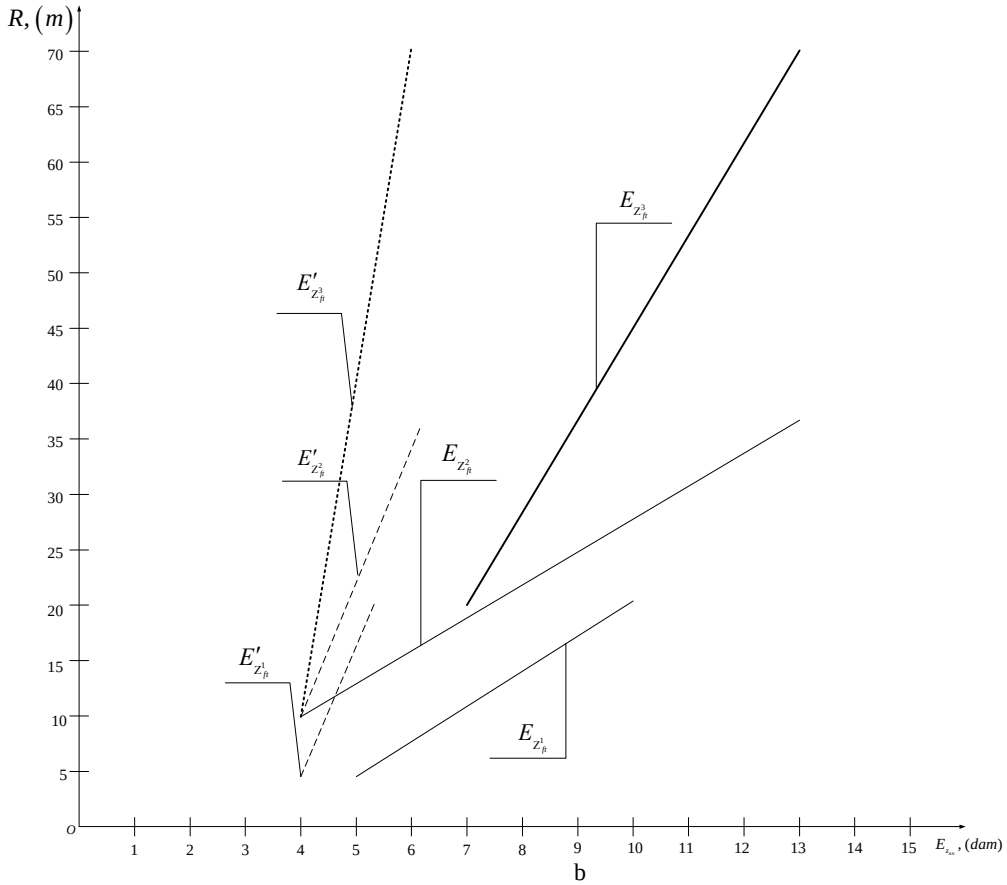
Table 3 – Total average errors in range and direction when determining the firing settings by the full preparation method, using a system of differential equations with and without coefficients

| Artillery system | Range, km | Median errors, м | | | |
|---|---|---|---|---|---|
| | | with coefficients | | without coefficients | |
| | | $E_{X\,fp}$ | $E_{Z\,fp}$ | $E_{X\,fp}$ | $E_{Z\,fp}$ |
| 122-mm BM-21 | 5 | 73 | 0–05 | 43 | 0–04 |
| | 10 | 141 | 0–06 | 85 | 0–04 |
| | 15 | 215 | 0–08 | 121 | 0–05 |
| | 20 | 283 | 0–10 | 165 | 0–05 |
| 220-mm 9P140 | 10 | 81 | 0–04 | 67 | 0–04 |
| | 20 | 262 | 0–07 | 167 | 0–05 |
| | 30 | 425 | 0–11 | 251 | 0–06 |
| | 36 | 505 | 0–13 | 308 | 0–06 |
| 300-mm 9A52 | 20 | 276 | 0–07 | 158 | 0–04 |
| | 30 | 433 | 0–08 | 249 | 0–04 |
| | 40 | 564 | 0–09 | 335 | 0–05 |
| | 50 | 717 | 0–10 | 425 | 0–05 |
| | 60 | 904 | 0–11 | 509 | 0–06 |
| | 70 | 980 | 0–13 | 595 | 0–06 |


DOI 10.15588/1607-3274-2024-4-19




a




b

Figure 2 – Total errors of the full preparation
a – by range for: 122-mm combat vehicle FM-21 –9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile and 300-mm 9A52 9M55C fragmentation projectile
b – by direction for: 122-mm combat vehicle FM-21 –9M21HE projectile, 220-mm combat vehicle 9P140 – 9M27HE projectile and 300-mm 9A52 9M55C fragmentation projectile


DOI 10.15588/1607-3274-2024-4-19

The system of equations (4) can be used for accurate calculation of the trajectory elements taking into account the wind coefficients of direct and cross wind effects on rockets separately for the active and passive sections of the trajectory and for the section of the flight trajectory of combat elements in the mathematical model of the method of complete preparation of the determination of firing units of the information-calculating component of the automated fire control system of combat vehicles of reactive artillery.

The dependencies of the calculation of the trajectory elements using a system of differential equations of the mathematical model of the method of full preparation of the determination of firing units of the information-calculation component of the automated fire control system of combat vehicles of jet artillery allow solving the problems of external ballistics with a perspective optimization of the trajectory of the projectile, which is used in reactive systems of salvo fire, as well as to determine a set of indicators that characterize the process of projectile flight approaching the tabular trajectory [5] and specify the initial conditions that are necessary to solve the system of equations of motion of the projectile on the active and passive sections of the trajectory and on the section of the flight trajectory of combat elements, which are considered as independent trajectories.

As the calculations showed, with the help of differential equations of the mathematical model of the method of complete preparation of the determination of firing installations of the information and calculation component of the automated fire control system of combat vehicles of reactive artillery (4), the accuracy of firing and the overall effectiveness of artillery fire depend on how successfully the task of ensuring the calculated movement of projectiles will be solved under the conditions of the influence of direct and cross, constant and variable wind on the trajectory.

## CONCLUSIONS

The improved mathematical model of the method of complete preparation of the determination of firing units of the information-calculation component of the automated fire control system of combat vehicles of reactive artillery allows to significantly speed up the process of preparing information for firing, and as a result ensures the promptness of response to changes in – a tribute to the software-mathematical complex of shooting and fire control.

It also increases the accuracy of their calculation at the same time.

The results of the calculated values show that the influence of various sources of error on the accuracy of the full training depended on the firing range, the artillery system, and the charge number.

The proposed system of differential equations of the mathematical model of the method of complete preparation of the determination of firing units of the information and calculation component of the automated system allows to take into account the weight functions of the air temperature, the influence of the wind for the active and passive sections of the trajectory of the projectile and the section of the detection of combat elements, to determine according to them, the weighting coefficients for each type of projectile, which makes it possible to quickly respond to changes in the tasks in the information-calculation process.

The system of differential equations of the mathematical model of the method for the complete preparation of the determination of firing units of the information and calculation component of the automated fire control system of combat vehicles of jet artillery can be used to calculate the values of corrections for the deviation of meteorological conditions of firing from the tabular conditions in fire divisions separately for each section of the missile's flight path, both existing and prospective.

Information of numerical calculations is resulted ground to assert:

1. Compared with the existing mathematical model of the method of full preparation for determining the settings for firing, the information-calculation component of the automated fire control system of combat vehicles of jet artillery will allow to shorten the equation when calculating the weighting coefficients for accounting for ballistic deviations of the subjects – air and ballistic wind conditions on active and passive sections of the trajectory and flight section of combat elements.

2. It will also ensure that when approximating the functions of the real distribution law, it is not replaced by a linear law, which will significantly increase efficiency and reduce the errors of calculating meteorological corrections by range and direction to 1.5–2%, will allow timely response to changing tasks in the information-calculation process during shooting and fire control, as well as effectively and quickly ensure the formation of formalized messages and documents based on the received information.

3. The improved mathematical model of the method of complete preparation of the determination of firing units of the information and calculation component of the automated fire control system of combat vehicles of jet artillery by solving the system of differential equations of motion of 9M21XE, 9M27XE and 9M55C projectiles allows hitting the specified targets with an accuracy of 0.5 – 0.9% in distance and 0–02 – 0–05 divisions of the protractor in direction and reduce the time for the preparation of installations.

## ACKNOWLEDGEMENTS

The authors of the article are grateful to the customer for the implementation of the research works "Harmonika", "Air", "Component" and for the scientific and applied justification of the direction of the research. The article corresponds to the field of technical science in the direction of information and computing processes of a component of an automated control system. The results


DOI 10.15588/1607-3274-2024-4-19

presented in the work were obtained and substantiated within the framework of the specified research works at the expense of state funding.

## REFERENCES

1. Celis R. D., Cadarso L., Sánchez J. Guidance and control for high dynamic rotating artillery rockets, *Aerospace Science and Technology*, 2017, Vol. 64, pp. 204–212. DOI:10.1016/j.ast. 2017.01.026.
2. Makyeyev V. I. Mathematical model spatial movement aircraft solid fuel in the atmosphere, *Messenger Sumy State University*, 2008, № 2, pp. 5–12.
3. Abbas L. K., Rui X. Numerical investigations of aero elastic divergence parameter of unguided launch vehicles, *Space Research Journal*, 2011, Issue 4(1), pp. 1–11. DOI: 10.3923/srj.2011.1.11
4. Burlov V., Lysenko L. Ballistics receiver systems. Moscow, Engineering, 2006, 459 p.
5. Morote J., Liaño G. Flight Dynamics of Unguided Rockets with Free-Rolling Wrap Around Tail Fins, *Journal of Spacecraft and Rockets*, 2006, Issue 43(6), pp. 1422–1423. DOI: 10.2514/1.22645
6. Sun H., Yu J., Zhang S. The control of asymmetric rolling missiles based on improved trajectory linearization control method, *Journal of Aerospace Technology and Management*, 2016, Vol. 8, Issue 3, pp. 319–327. DOI: 10.5028/jactm.v813.617
7. Arutyunova N. K., Dulliev A. M., Zabotin V. I. Models and methods for three external ballistics inverse problems, *Bulletin of the South Ural State University. Ser. Mathematical Modelling, Programming & Computer Software (Bulletin SUSU MMCS)*, 2017, Vol. 10, Issue 4, pp. 78–91. DOI: 10.14529/mmp170408
8. Lahti J. , Sailaranta T., Harju M., Virtanen K. Control of exterior ballistic properties of spin-stabilized bullet by optimizing internal mass distribution, *Defence Technology*, 2019, Vol. 15, pp. 38–50. DOI: 10.1016/j.dt.2018.10.003
9. Kokes J., Costello M., Sahu J. Generating an aerodynamic model for projectile flight simulation using unsteady time accurate computational fluid dynamic results, *WIT Transactions on Modelling and Simulation*, 2007, Vol. 45, pp. 31–54. DOI: 10.2495/CBAL070041
10. McCoy R. Modern Exterior Ballistics: The Launch and Flight Dynamics of Symmetric Projectiles. Pennsylvania: Schiffer Publishing Ltd., 2012, 328 p.
11. Makeev V. I., Grabchak V. I., Trofimenko P. E., Pushkarev Y. I. Research of jet engine operation parameters on the range and accuracy of firing of rockets, *Information processing system*, 2008, Vol. 6(73), pp. 77–81.
12. Grabchak V. I., Makeev V. I., Trofimenko P. E., Pushkarev Yu. I. Substantiation of a rational correction system for firing with active rockets (mines), *Artillery and small arms*, 2009, Vol. 4, pp. 3–9.
13. Majstrenko O. V., Prokopenko V. V., Makeev V. I., Ivanyk E. G. Analytical methods of calculation of powered and passive trajectory of reactive and rocket-assisted projectiles, *Radio Electronics, Computer Science, Control Journal*, 2020, Issue 2(53), pp. 173–182. DOI 10.15588/1607-3274-2020-2-18.
14. Lahti J., Sailaranta T., Harju M., Virtanen K. Control of exterior ballistic properties of spin-stabilized bullet by optimizing internal mass distribution, *Defense Technology*, 2019, Vol. 15, pp. 38–50. DOI: 10.1016/j.dt.2018.10.003Received 00.00.2020.
15. Makeiev V. I., Lapa M. M, Latin S. G., Trofimenko P. E. Methods for determining corrections for nonlinearity and interaction of perturbing factors, *National Academy of Sciences of Ukraine Institute of Modem ing Problems in Energy named after G. E. Puhov. Electronic modelling*, 2012, Issue 34( 1), pp. 109–119.
16. Gao F., Zhang H. Study of 2-D trajectory correction based on geomagnetic detection with impulse force for projectiles, *Journal of System Simulation*, 2011, Vol. 23, pp. 123-128.
17. Xiu-Ling J. L, Wang H. P., Zeng S.M. et al. CFD prediction of longitudinal aerodynamics for a spinning projectile with fixed canard, *Transactions of Beijing Institute of Technology*, 2011, Vol. 31, pp. 265–268.
18. Korn G. A., Korn T. M. Mathematical Handbook for Scientists and Engineers. New York, McGraw-Hill Book Company, 1968, 832 p.
19. Dmitrievsky A. A., Lysenko L.N. External ballistics: a textbook for university students. Moscow, Engineering, 2005, 608 p.

Received 29.01.2024.
Accepted 05.11.2024.


DOI 10.15588/1607-3274-2024-4-19

УДК 629.072.19

# УДОСКОНАЛЕНА МАТЕМАТИЧНА МОДЕЛЬ СПОСОБУ ПОВНОЇ ПІДГОТОВКИ ВИЗНАЧЕННЯ УСТАНОВОК ДЛЯ СТРІЛЬБИ НА УРАЖЕННЯ ІНФОРМАЦІЙНО-РОЗРАХУНКОВОЇ СКЛАДОВОЇ АВТОМАТИЗОВАНОЇ СИСТЕМИ УПРАВЛІННЯ ВОГНЕМ БОЙОВИХ МАШИН РЕАКТИВНОЇ АРТИЛЕРІЇ

**Майстренко О. В.** – д-р військових наук, провідний науковий співробітник Науково-методичного центру організації наукової та науково-технічної діяльності Національного університету оборони України, полковник, Київ, Україна.

**Макєєв В. І.** – канд. технічних наук, доцент, доцент військового відділення Сумського державного університету, полковник, Суми, Україна.

**Прокопенко В. В.** – канд. техн. наук, заступник начальника науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, підполковник, Львів, Україна.

**Андреєв І. М.** – старший науковий співробітник науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, полковник, Львів, Україна.

**Каменцев С. Ю.** – провідний науковий співробітник науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, полковник, Львів, Україна.

**Онофрійчук А. Я.** – молодший науковий співробітник науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, майор, Львів, Україна.

## АНОТАЦІЯ

**Актуальність.** У рамках автоматизації системи управління вогнем бойових машин реактивної артилерії, щодо підготовки даних для стрільби та управління вогнем, удосконалено інформаційно-обчислювальний процес цієї системи, а саме удосконалено математичну модель способу повної підготовки визначення установок для стрільби снарядами які застосовуються в реактивних системах залпового вогню. В систему диференційних рівнянь математичної моделі інформаційно-обчислювального процесу складової автоматизованої системи управління вогнем бойових машин реактивної артилерії, введено вагові функції по температурі повітря, впливу вітру для активної, пасивної ділянок траєкторії польоту снаряду та ділянки розкриття бойових елементів, що дозволяє визначати по ним вагові коефіцієнти для кожного типу снаряду.

**Мета роботи.** Удосконалити інформаційно-розрахункову складову автоматизованої системи управління вогнем бойових машин реактивної артилерії, шляхом покращення математичної моделі способу повної підготовки визначення установок для стрільби на ураження. Запропонувавши систему диференційних рівнянь яка буде враховувати вагові функції по температурі повітря, впливу вітру для активної, пасивної ділянок траєкторії польоту снаряду та ділянки розкриття бойових елементів, а також дасть можливість визначати по ним вагові коефіцієнти для кожного типу снаряду, що в свою чергу призведе до підвищення точності визначення установок для стрільби.

**Метод.** Запропонований аналітичний метод дозволяє: провести розрахунок вагових коефіцієнтів для кожного типу реактивного снаряду, що характеризують процес наближення польоту реактивного снаряду до табличної траєкторії та задати початкові умови необхідні для вирішення диференціальних рівнянь математичної моделі інформаційно-обчислювального процесу складової автоматизованої системи управління вогнем бойових машин реактивної артилерії; підвищити точності визначення установок для стрільби при виконанні вогневих завдань, що надає змогу оперативно реагувати на зміну бойової обстановки шляхом змін у програмно-математичному процесі автоматизованої системи управління вогнем; ефективно та оперативно забезпечити розробку або уточнення текстуальних і графічних розпорядчих та бойових документів на основі отриманих результатів за допомогою диференціальних рівнянь математичної моделі інформаційно-обчислювального процесу складової автоматизованої системи управління вогнем.

**Результати.** Система диференціальних рівнянь математичної моделі інформаційно-обчислювального процесу складової автоматизованої системи управління вогнем бойових машин реактивної артилерії забезпечує своєчасне реагування на зміну обстановки в інформаційно-обчислювальному процесі складової автоматизованої системи управління вогнем бойових машин реактивної артилерії під час стрільби та управління вогнем. Надає можливість ефективно та оперативно забезпечити розроблення або уточнення текстуальних і графічних розпорядчих та бойових документів за інформацією, отриманою під час виконання вогневих завдань.

**Висновки.** Проведені розрахунки на основі запропонованої системи диференційних рівнянь підтверджують покращення інформаційно-розрахункової складової автоматизованої системи управління вогнем бойових машин реактивної артилерії та дозволяють своєчасно реагувати на зміну завдань в інформаційно-розрахунковому процесі під час стрільби і управління вогнем, а також ефективно і швидко забезпечують формування формалізованих повідомлень і документів на основі отриманої інформації під час виконання вогневого завдання підрозділами реактивної артилерії. Перспективами подальших досліджень є створення узгоджених математичних методів, моделей, алгоритмів і програм для реалізації цілей і завдань стрільби і управління вогнем при складанні Таблиць стрільби для перспективних або отриманих бойових машин реактивної артилерії від партнерів.

**КЛЮЧОВІ СЛОВА:** автоматизована система управління, інформаційно-розрахункова складова, математична модель, система диференційних рівнянь, апроксимація функцій реального закону розподілу, вагові функції по температурі повітря, впливу вітру.


DOI 10.15588/1607-3274-2024-4-19

## REFERENCES

1. Celis R. D. Guidance and control for high dynamic rotating artillery rockets / R. D Celis, L. Cadarso, J. Sánchez // Aerospace Science and Technology. – 2017. – Vol. 64. – P. 204–212. DOI:10.1016/j.ast. 2017.01.026.
2. Makyeyev V. I. Mathematical model spatial movement aircraft solid fuel in the atmosphere / V. I. Makyeyev // Messenger Sumy State University. – 2008. – № 2. – P. 5–12.
3. Abbas L. K. Numerical investigations of aero elastic divergence parameter of unguided launch vehicles / L. K. Abbas, X. Rui // Space Research Journal. – 2011. – Issue 4(1). – P. 1–11. DOI: 10.3923/srj.2011.1.11
4. Burlov V. Ballistics receiver systems / V. Burlov, L. Lysenko. – Moscow: Engineering, 2006. – 459 p.
5. Morote J. Flight Dynamics of Unguided Rockets with Free-Rolling Wrap Around Tail Fins / J. Morote, G. Liaño // Journal of Spacecraft and Rockets. – 2006. – Issue 43(6). – P. 1422–1423. DOI: 10.2514/1.22645
6. Sun H. The control of asymmetric rolling missiles based on improved trajectory linearization control method / H. Sun, J. Yu, S. Zhang // Journal of Aerospace Technology and Management. – 2016. – Vol. 8, Issue 3. – P. 319–327. DOI: 10.5028/jactm.v813.617
7. Arutyunova N. K. Models and methods for three external ballistics inverse problems / N. K. Arutyunova, A. M. Dulliev, V. I. Zabotin // Bulletin of the South Ural State University. Ser. Mathematical Modelling, Programming & Computer Software (Bulletin SUSU MMCS). – 2017. – Vol. 10, Issue 4. – P. 78–91. DOI: 10.14529/mmp170408
8. Control of exterior ballistic properties of spin-stabilized bullet by optimizing internal mass distribution / [J. Lahti, T. Sailaranta, M. Harju, K. Virtanen] // Defence Technology. – 2019. – Vol. 15. – P. 38–50. DOI: 10.1016/j.dt.2018.10.003
9. Kokes J. Generating an aerodynamic model for projectile flight simulation using unsteady time accurate computational fluid dynamic results / J. Kokes, M. Costello, J. Sahu // WIT Transactions on Modelling and Simulation. – 2007. – Vol. 45. – P. 31–54. DOI: 10.2495/CBAL070041
10. McCoy R. Modern Exterior Ballistics: The Launch and Flight Dynamics of Symmetric Projectiles / R. McCoy. – Pennsylvania: Schiffer Publishing Ltd., 2012. – 328 p.
11. Research of jet engine operation parameters on the range and accuracy of firing of rockets / [V. I. Makeev, V. I. Grabchak, P. E. Trofimenko, Y. I. Pushkarev] // Information processing system. – 2008. – Vol. 6(73). – P. 77–81.
12. Substantiation of a rational correction system for firing with active rockets (mines) / [V. I. Grabchak, V. I. Makeev, P. E. Trofimenko, Yu. I. Pushkarev] // Artillery and small arms. – 2009. – Vol. 4. – P. 3–9.
13. Analytical methods of calculation of powered and passive trajectory of reactive and rocket-assisted projectiles / [O. V. Majstrenko, V. V. Prokopenko, V. I. Makeev, E. G. Ivanyk] // Radio Electronics, Computer Science, Control Journal. – 2020. – Issue 2(53). – P. 173–182. DOI 10.15588/1607-3274-2020-2-18.
14. Control of exterior ballistic properties of spin-stabilized bullet by optimizing internal mass distribution / [J. Lahti, T. Sailaranta, M. Harju, K. Virtanen] // Defense Technology. – 2019. – Vol. 15. – P. 38–50. DOI: 10.1016/j.dt.2018.10.003Received 00.00.2020.
15. Methods for determining corrections for nonlinearity and interaction of perturbing factors / [V. I. Makeiev, M. M. Lapa, S. G. Latin, P. E. Trofimenko] // National Academy of Sciences of Ukraine Institute of Modem ing Problems in Energy named after G. E. Puhov. Electronic modelling. – 2012. – Issue 34( 1). – P. 109–119.
16. Gao F. Study of 2-D trajectory correction based on geomagnetic detection with impulse force for projectiles / F. Gao, H. Zhang // Journal of System Simulation. – 2011. – Vol. 23. – P. 123–128.
17. CFD prediction of longitudinal aerodynamics for a spinning projectile with fixed canard / [J. L. Xiu-Ling, H. P. Wang, S. M. Zeng et al.] // Transactions of Beijing Institute of Technology. – 2011. – Vol. 31. – P. 265–268.
18. Korn G. A. Mathematical Handbook for Scientists and Engineers / G. A. Korn, T. M. Korn. – New York : McGraw-Hill Book Company, 1968. – 832 p.
19. Dmitrievsky A. A. External ballistics: a textbook for university students / A. A. Dmitrievsky, L. N. Lysenko. – Moscow : Engineering, 2005. – 608 p.


DOI 10.15588/1607-3274-2024-4-19

UDC 622.7: 534

# GENERAL PRINCIPLES OF FORMALIZATION OF TECHNOLOGICAL PROCESS CONTROL OF MINING PRODUCTION IN A DYNAMIC DISTRIBUTED SYSTEM

**Morkun V. S. –** Dr. Sc., Professor, Professor at the University of Bayreuth, Bayreuth, Germany.
**Morkun N. V. –** Dr. Sc., Professor, Professor at the University of Bayreuth, Bayreuth, Germany.
**Hryshchenko S. M.** – PhD, Senior researcher in the specialty Automation and computer-integrated technologies State Tax University, Irpin, Ukraine.
**Shashkina A. A. –** Postgraduate student at Kryvyi Rih National University, Kryvyi Rih, Ukraine.
**Bobrov E. Y. –** Postgraduate student at Kryvyi Rih National University, Kryvyi Rih, Ukraine.

## ABSTRACT

**Context.** The problem of synthesis, modeling, and analysis of automated control of complex technological processes of mining production as a dynamic structure with distributed parameters.

**Objective**. On the example of the technological line of ore beneficiation, the general principles of formalization of control of mining production processes as a dynamic system with distributed parameters are considered.

**Method.** The modeling of interactions between individual components of the control system is carried out using the methods of coordinated distributed control. In accordance with this approach, the technological line is decomposed into a set of separate subsystems (technological units, enrichment cycles). Under these circumstances, the solution to the global optimization problem is also decomposed into a corresponding set of individual subproblems of optimizing the control of subsystems. To solve the global problem, this formulation uses a two-level structure with coordinating variables that are fed to the input of local control systems for technological units and cycles. At the lower level of control, sets of subtasks have independent solutions, coordinated by the coordinating variables formed at the upper level.

**Results.** The paper proposes a method for forming control of a distributed system of technological units of an ore dressing line based on the decomposition of the dynamics of the distributed system into time and space components. In the spatial domain, the control synthesis problem is solved as a sequence of approximation problems of a set of spatial components of the dynamics of the controlled system. In the time domain, the solution of the control synthesis problem is based on the methods of synthesizing control systems with concentrated parameters.

**Conclusions.** The use of the proposed approach to the formation of technological process management at mining enterprises of the Kryvyi Rih iron ore basin will improve the quality of iron ore concentrate supplied to metallurgical processing, increase the productivity of technological units and reduce energy consumption.

**KEYWORDS:** mining, automation, ore dressing, distributed control, process, system.

## ABBREVIATIONS

DCSs is a design and implementation of distributed control systems;

ESRD is a system with Concentrated Input and Distributed Output Variables with Extrapolator; CCD is a system with Concentrated Input and Distributed Output Variable.

## NOMENCLATURE

$Q_n$ is a productivity of stages for the finished product;

$Q_i^3$ is a set performance values;

$\beta_n$ is a content of a useful component in the product of the stage;

$\beta_i^{\min}$ is a limits of useful component content by stage;

$\beta_{XB1}$ is a losses of useful component in the tailings by stages and in general at the outlet of the enrichment line;

$\beta_{XB1}^{\min}$ is a restrictions on tail losses;

$\xi$ is a specific gravity for each technological type in ore;

$\Omega$ is a current, minimum and maximum amount of concentrate produced, respectively;

$\beta$ is a current, minimum and maximum proportion of total iron in the concentrate, respectively;

$E$ is a mass of ore that can be processed by the concentrator;

$Q$ is a total volume of ore processing of all technological types;

$\beta^d(l,s)$ is an output variable of this system is equivalent to the output variable of a system with distributed input and distributed output;

$Q(\overline{p})$ is a productivity of the ore dressing line of finished product;

$\beta(\overline{p})$ is a content of a useful component in the finished product;

$\overline{p}$ is a vector of coordinating parameters of the second level of control, which form the conditions for solving the tasks of the first level;


DOI 10.15588/1607-3274-2024-4-20

$\Theta H_i(\ell_i,k)$ is a distributed impulse function of the system;

$\Theta H_i(\ell,(k+\varepsilon)T)$ is a distributed impulse response transient function of a ESRD, consisting of a CCD and a zero-order extrapolator unit Hi;

$\Theta H_i(\ell_i,t)$ is a partially distributed output variable in the time domain at a point $\ell_i$;

$\{\psi_i(k)\}_{i,k}$ is a sequences of input variables ESRD;

$\{\Theta H_i(\ell_j,k)\}_{i,j,k}$ are the values of discrete distributed impulse response functions;

$q$ is a discrete transition time CCD;

$\Omega HR_i(\ell,\infty)$ is a constant values of the given transient characteristics of the distributed parameters of the technological line of beneficiation;

$B_i{}^d(\ell_i,\infty)$ are the approximation parameters;

$\breve{\breve{B}}^d$ is a vector of optimal approximation;

$\left(\beta_i^d\right)'(\ell,k)$ is an approximation of the distributed output variables of the identified object;

$\Xi_i(\ell_i,s)$ is an appropriate transfer functions;

$K_i(z)$ is a robust regulators;

$\Omega HR_i(\ell,\infty)$ is a constant values of the given transient characteristics of the distributed parameters of the technological line of beneficiation;

$B_i{}^d(\ell_i,\infty)$ are the approximation parameters;

$\breve{\breve{B}}^d=\left\{\breve{B}_i{}^d(\ell_i,\infty)\,|\,i=\overline{1,n}\right\}$ is a vector of optimal approximation;

$\left(\beta_i^d\right)'(\ell,k)$ is an approximation of the distributed output variables of the identified object.

## INTRODUCTION

The technological processes of mining production, namely the extraction and preparation of raw materials for metallurgical processing, are distributed in space, characterized by various dynamics, many interrelated variables and disturbing factors. An example of a corresponding diagram of the chain of devices built on the basis of a typical technological line of an ore processing plant with the indication of control points for the characteristics of ore material at different stages of iron ore processing is shown in Fig. 1.

A characteristic feature of technological complexes of concentration plants as control objects is the multistage nature of production processes implemented using many units, equipment, and transport links between them through a multi-circuit water-sludge circuit and/or belt conveyors. At the same time, positive technological feedback (re-cycle) is organized through the water-sludge circuit, both within and between the beneficiation stages.

Synthesis, modeling, and optimizati on of the management of such mining structures is still a problematic issue.

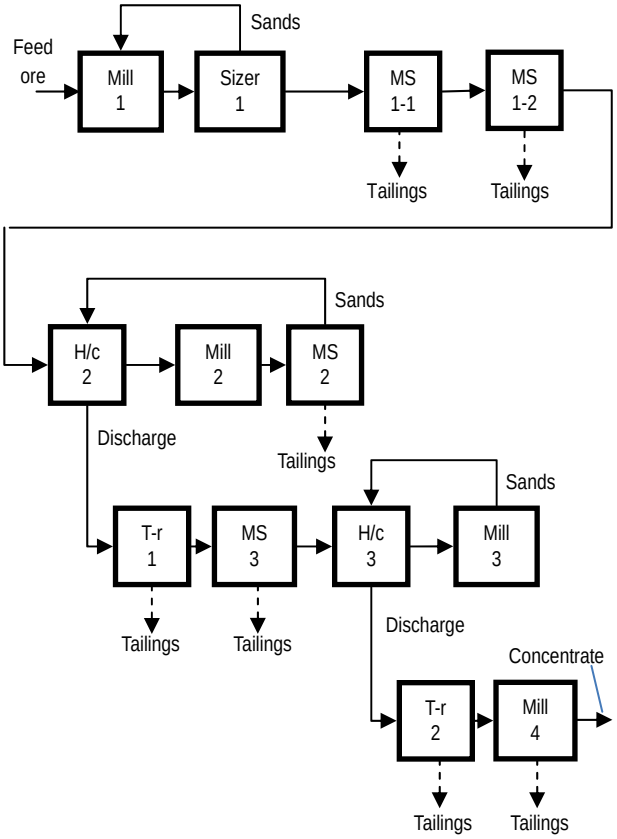



Figure 1 – Diagram of the chain of devices on the technological line of the ore dressing plant

**The object of study** is processes of automated control of the technological line of mining production in the conditions of changing characteristics of raw materials and the state of technological units/

**The subject of study** is the mathematical models, criteria and methods of distributed optimal control of interrelated processes of mining production based on a dynamic spatial and temporal structure

**The purpose of the work** of the research is to improve the energy efficiency and quality of automated control of the technological line for the preparation of ore for metallurgical processing, to increase the extraction of useful components into concentrate during the processing of iron ore by developing principles and approaches to distributed optimal control of interrelated mining processes based on a dynamic space-time model.

## 1 PROBLEM STATEMENT

Multi-criteria control structures formed on the basis of a complex indicator that includes the following goals are widely used in the automated control systems of concentrating production: maximizing ore processing productivity, maximizing concentrate quality, and minimizing losses of useful components in tailings [1].


DOI 10.15588/1607-3274-2024-4-20

$$J = \begin{cases} Q \to \max; \\ \beta_k \to \max; \\ \beta_{\text{XB}} \to \min. \end{cases} \quad (1)$$

An expanded version of criterion (1), which comprehensively takes into account the performance indicators of individual stages of ore dressing, is proposed in [2].

$$J = \begin{cases} Q_1 \to \max(Q_1 \to Q_1^3);\ \beta_1^{\min} \le \beta_1 \le \beta_1^{\max};\ \beta_{XB1}^{\min} \le \beta_{XB1} \le \beta_{XB1}^{\max}; \\ Q_2 \to \max(Q_2 \to Q_2^3);\ \beta_2^{\min} \le \beta_2 \le \beta_2^{\max};\ \beta_{XB2}^{\min} \le \beta_{XB2} \le \beta_{XB2}^{\max}; \\ \ldots \\ Q_K \to \max(Q_K \to Q_K^3);\ \beta_K^{\min} \le \beta_K \le \beta_K^{\max};\ \beta_{XB}^{\min} \le \beta_{XB} \le \beta_{XB}^{\max}. \end{cases} \quad (2)$$

In the existing automated process control systems in the mining and metallurgical complex, one of the most important indicators along with technological indicators is the energy efficiency indicator [3–6]. Increasing the energy efficiency of the control of the technological process of beneficiation, as shown in [7], is achieved by applying a comprehensive criterion

$$\begin{cases} \left|\overline{\xi} - \overline{\xi^*}\right| \to \min; \\ \Omega_l \le \Omega \le \Omega_h; \\ \beta_l \le \beta \le \beta_h; \\ Q/E \to \max. \end{cases} \quad (3)$$

This criterion ensures that the processed ore maintains the optimal particle size distribution in terms of recovery of the useful component, taking into account the mass ratio of certain technological types, and that the concentrate of a given quality is produced at the maximum utilization rate of technological processing equipment.

The analysis of control systems for the technological process of iron ore beneficiation showed that the control object should be represented as a distributed system [8, 9–11] with a concentrated input and a distributed output. At the same time, the output variable of this system is equivalent to the output variable of a system with distributed input and distributed output $\beta^d(l,s)$ [12].

Based on the above results, it is advisable to use a global optimization criterion when forming the control of an ore dressing line as a dynamic system with distributed parameters, which is generally described by the system

$$\begin{cases} Q(\overline{p}) \to \max; \\ \beta_l \le \beta(\overline{p}) \le \beta_h. \end{cases} \quad (4)$$

## 2 REVIEW OF THE LITERATURE

The DCSs are associated with many challenging problems, including network-induced delays, time-varying topology or throughput, or increasing complexity [12]. The need to solve these problems has stimulated the development of methods for modeling, analyzing, and synthesizing RSCs [13, 14].

Paper [13] presents the proposed methodology for modeling and analyzing distributed control systems, consisting of three modules, each of which is divided into several submodules. The methodology includes steps for analyzing communication interfaces and computational processes, and also presents a method for creating models of DCSs components in the form of colored time Petri nets and finite state machines. The result of the applied methodology is the analysis of the DSC, such as prediction of the throughput and response of the system to various input factors.

Computational processes and communication networks in DCSs can be viewed as systems with discrete events, which can be conveniently modeled using a number of approaches, including finite state machines and Petri nets [15–17].

In [18], the study of distributed control systems is considered from the perspective of “network management”, providing a snapshot of five control issues: sampling control, quantization control, network control, event-driven control, and security control.

A DCSs is characterized by a multi-level architecture consisting of numerous subsystems deployed in a distributed manner, where control functions are distributed among many controllers, and individual control levels are connected by different types of communication networks [19].

In many cases, these subsystems have three fundamental characteristics [20]. The first characteristic is autonomy. Subsystems are usually autonomous or semi-autonomous and can often exist and operate independently. The second characteristic is homogeneity. Subsystems regularly have similarities and play relatively equal roles. The last and most important characteristic is interactivity. Subsystems must be connected to the system topology collectively so that they can communicate by exchanging information or substance and cooperate with each other.

Paper [21] analyzes the functionality, requirements, and cost of installation and operation of industrial DCSs systems based on PLC logic controllers. Particular attention is paid to the distribution of individual system components by the corresponding layers contained in their structure.

Distributed control systems offer numerous advantages over traditional centralized control systems. The main advantages of DCSs usually include the following properties [22]:

– increased reliability: by distributing control functions among several controllers, the DCSs eliminates single points of failure;

– scalability: DCSs makes it easy to expand control systems as industrial operations grow;

– flexibility: The DCSs provides full integration with other control systems, such as PLCs and SCADA systems;


DOI 10.15588/1607-3274-2024-4-20

– reduced operating costs: By automating complex processes, the DCSs can help industries reduce labor costs, minimize human error, and optimize resource utilization;

– security: cybersecurity and physical security of the system are ensured at the operator and engineer levels.

But, like other advanced technologies, the implementation of DCSs also has certain problems and limitations [22]:

– high initial investment: the implementation of a DCSs may involve the purchase and installation of many controllers, I/O modules, communication networks, and human-machine interfaces;

– complexity: designing and configuring a DCSs requires specialized knowledge and experience;

– maintenance and support: Like any complex system, a DCSs requires regular maintenance and troubleshooting to ensure optimal performance.

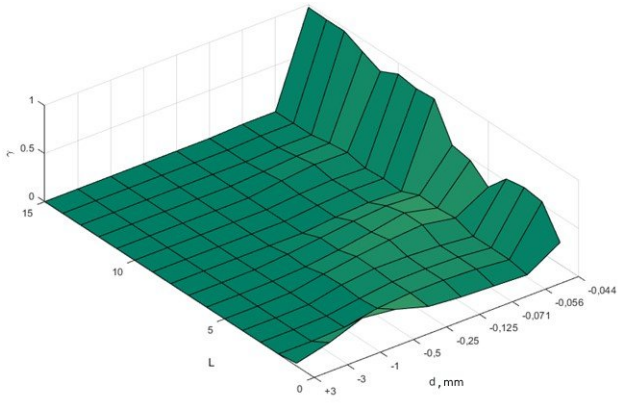
In the mining and ore processing industry, DCSs are effectively used to automate technological operations for reducing ore size [23–25], classifying grinding products [26–28], underground ore mining [29, 30], etc. The application of this approach is well coordinated with the systems of non-destructive non-contact control of the characteristics of raw materials and processed products as a means of information support for control processes [28, 31, 32].

The results of the analysis of studies on the use of distributed control systems for complex technological processes show that their use helps to increase production efficiency by providing centralized monitoring and control of distributed processes, allowing operators to make changes in real time to optimize performance [22]. In addition, RSCs use advanced control algorithms and data analytics to optimize processes, reduce energy consumption, and minimize waste.

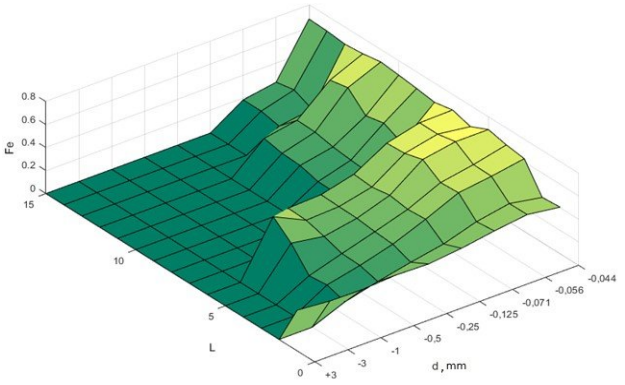
## 3 MATERIALS AND METHODS

The characteristics of iron ore raw materials in the process of multistage processing are presented as spatial and temporal variables distributed over an interval. The function of the distribution of the content of the useful component in the particle size classes of the ore material being processed is taken as the initial distributed variable [4].

After applying concentrated control actions to the system input, we obtain the system output signal as a distributed variable $\beta_i^d(\ell,t)$. An example of the realization of the output signal of the control object is, in this case, the output of size classes (Fig. 2a) and the iron content in size classes (Fig. 2b), distributed over the technological operations of the ore dressing line.




a




b

Figure 2 – Distributed along the technological line of beneficiation:
a) yield of solid of the pulp phase;
b) indicator of iron content content in the industrial product

The relationship between output and input variables is described by convolution [11].

$$\beta_i^d(\ell,t)=\Theta_i(\ell,t)*\psi_i(t). \quad (5)$$

In the event that $\psi_i(t)$ is a single step function at the output of the system, we obtain a distributed transition function $\Omega_i(\ell,t)$. Similarly, we obtain the relationship between the individual input variables $\{\psi_i(t)\,|\,i=\overline{1,n}\}$ and the corresponding distributed output variables $\{\beta_i^d(\ell,t)\,|\,i=\overline{1,n}\}$. The total distributed function of the distribution of the content of the corrosive component by size classes is described by the expression

$$\beta^d(\ell,t)=\sum_{i=1}^{n}\Theta_i(\ell,t)*\psi_i(t). \quad (6)$$

To increase the accuracy of identification of model parameters, it is advisable to measure the characteristics of ore material at the outlet of technological aggregates, which in the considered system correspond to points with


DOI 10.15588/1607-3274-2024-4-20

coordinates $\ell_1,\ldots,\ell_n$, located near the points of application of controlling influences $\psi_1(t),\ldots,\psi_n(t)$, which can be, in particular [3, 4]: ore and water consumption in the mill, water consumption in the classifier, technological sump, hydrocyclone, magnetic separator, deslimer, etc. At these points, the output variables are partially distributed $\beta_1^d(\ell_1,t),\ldots,\beta_n^d(\ell_n,t)$, located on the respective surfaces of the distributed output variables $\beta_1^d(\ell,t),\ldots,\beta_n^d(\ell,t)$. Also, at the following locations $\ell_1,\ldots,\ell_n$ partial distributed impulse response functions are determined on the surfaces of individual distributed impulse response functions $\Theta_1(\ell_1,t),\ldots,\Theta_n(\ell_n,t)$. Using the relation (5), we obtain

$$\beta_i^d(\ell_i,t)=\Theta_i(\ell_i,t)*\psi_i(t),i=\overline{1,n}. \quad (7)$$

Taking into account the discrete input variables and the zero-order extrapolator {Hi} and the discrete outputs of the model of the ore dressing technological line are represented as a convolution

$$\beta_i^d(\ell,(k+\varepsilon)T)=\Theta H_i(\ell,(k+\varepsilon)T)*\psi_i(kT). \quad (8)$$

Let us assume T=1, ε=0. Then, according to formula (8), we obtain an expression for the function of the distribution of the useful component by fineness classes at the output of the ore dressing line

$$\beta_i^d(\ell,k)=\Theta H_i(\ell,k)*\psi_i(k),\quad i=\overline{1,n}. \quad (9)$$

Common distributed output variable $\beta^d(\ell,k)$ of the processing line as a control object with a concentrated input and distributed output with zero-order extrapolatorsy $\{H_i\}_i$, and will be written as

$$\beta^d(\ell,k)=\sum_{i=1}^{n}\Theta H_i(\ell,k)*\psi_i(k). \quad (10)$$

For points $\ell_1,\ldots,\ell_n$ located around concentrated input variables, partially distributed output variables can be obtained $\beta_1^d(\ell_1,k),\ldots,\beta_n^d(\ell_n,k)$ and consequently, partially distributed impulse transition functions $\Theta H_1(\ell_1,k),\ldots,\Theta H_n(\ell_n,k)$. In this case, the following convolutions are correct

$$\beta_i^d(\ell_i,k)=\Theta H_i(\ell_i,k)*\psi_i(k),\quad i=\overline{1,n}. \quad (11)$$

After applying the Laplace transform, moving to the domains, we obtain

$$\beta_i^d(\ell,s)=\Xi_i(\ell,s)\psi_i(s),\ i=\overline{1,n}, \quad (12)$$

$$\beta^d(\ell,s)=\sum_{i=1}^{n}\Xi_i(\ell,s)\psi_i(s), \quad (13)$$

$$\beta_i^d(\ell_i,s)=\Xi_i(\ell_i,s)\psi_i(s),\ i=\overline{1,n}. \quad (14)$$

In the case when the action of the elements of the vector $\{\psi_i(t)\,|\,i=\overline{1,n}\}$ is added at points $\ell_1,\ldots,\ell_n$ at the outputs of the units $\Xi H_1(\ell_1,s),\ldots,\Xi H_n(\ell_n,s)$ Partially distributed output variables are formed $\beta_1^d(\ell_1,k),\ldots,\beta_n^d(\ell_n,k)$. Distributed output variables of blocks $\Xi H_1(\ell,s),\ldots,\Xi H_n(\ell,s)$: $\beta_1^d(\ell,k),\ldots,\beta_n^d(\ell,k)$ move along these trajectories. Thus, the total output variable of the technological line of ore dressing, as a distributed object, is given by the ratio:

$$\beta^d(\ell,k)=\sum_{i=1}^{n}\beta_i^d(\ell,k). \quad (15)$$

When a discrete representation is considered, the discrete dynamics of the system is given by a set of discrete distributed impulse response functions $\{\Theta H_i(\ell,k)\,|\,i=\overline{1,n}\}$. The given characteristics on this set are written as follows

$$\{\Theta HR_i(\ell,k)=\Theta H_i(\ell,k)/\Theta H_i(\ell_i,k)\,|\,i=\overline{1,n}\}. \quad (16)$$

This makes it possible to determine the output of the distributed control system of the processing line as follows

$$\beta_i^d(\ell,k)=\Theta H_i(\ell_i,k)\Theta HR_i(\ell,k)*\psi_i(k),\quad i=\overline{1,n}\quad i=\overline{1,n}. \quad (17)$$

The dynamics of the technological line of ore dressing, as a distributed object, is decomposed into a time component that depends on the continuous (*t*) or discrete (*k*) time at fixed values of *i*, $\ell_i:\{\Theta H_i(\ell_i,t)\}_{i,k}$, and the spatial component, which depends on a simple variable $\ell_i$ at fixed values of *i*, *k*: $\{\Theta HR_i(\ell,t)\}_{i,k}$.

Steady state of the function of distribution of the useful component by ore material size classes as distributed along the technological line of beneficiation of the initial $\beta^d(\ell,\infty)$ variable, in accordance with the theory proposed in [8, 12], can be described by means of a stable value of the transition function, $\{\Omega H_i(\ell,\infty)\,|\,i=\overline{1,n}\}$, that is obtained after applying a single stepwise impact 1(*k*) to each input $\{\psi_i(k)\,|\,k=\overline{1,n}\}$.

When forming the control, we use the following decomposition of the distributed dynamics of the system:


DOI 10.15588/1607-3274-2024-4-20

time component $\left\{\Xi H_i(\ell_i,s)\,|\,i=\overline{1,n}\right\}$, spatial component $\left\{\Omega HR_i(\ell,\infty)\,|\,i=\overline{1,n}\right\}$.

As a result of spatial discretization of the model with discrete time of the ore dressing line, which has concentrated inputs (adjustable parameters of technological units and operations) and distributed output (content of the useful component classes) at the points $\left\{\ell_j \mid j=\overline{1,m}\right\}$ is obtained in scalar form

$$\left\{\beta^d(\ell_i,k)=\sum_{i=1}^{n}\Theta H_i(\ell_j,k)*\psi_i(k)\right\}_{j,k,i}. \qquad (18)$$

Expression (14) is written in vector form as follows

$$\bar{\beta}^d(\ell_i,k)=\sum_{i=1}^{n}\Theta H_i(\ell_j,k)*\psi_i(k). \qquad (19)$$

Considering that for a particular $i, j, r > q$ The following ratio is true: $\Theta H_i(\ell_j,r)=0$, where $q$ – discrete transition time CCD, expression (19) can be written as

$$\begin{bmatrix}\beta^d(\ell_1,k)\\ \dots\\ \beta^d(\ell_m,k)\end{bmatrix}=\sum_{i=1}^{n}\begin{bmatrix}\Theta H_i(\ell_1,0) & \dots & \Theta H_i(\ell_1,q)\\ \dots & \dots & \dots\\ \Theta H_i(\ell_m,0) & \dots & \Theta H_i(\ell_m,q)\end{bmatrix}\begin{bmatrix}\psi_i(k)\\ \dots\\ \psi_i(k-q)\end{bmatrix}. \qquad (20)$$

In expression (16), the individual matrices correspond to the distributed discrete impulse transition functions of the considered distributed object of controlling the ore dressing process, which is represented by different mineralogical and technological varieties (Fig. 3): FDV – formers of distributed variables, EECV – executive elements of concentrated variables.

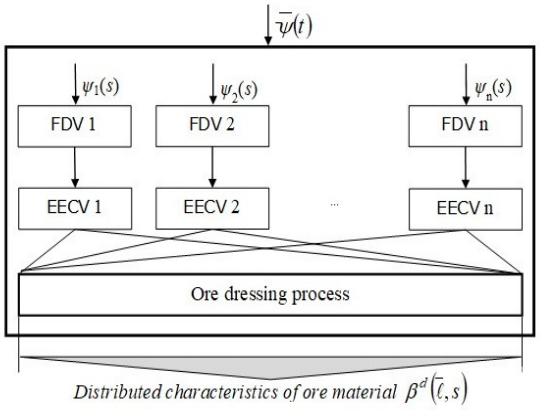



Figure 3 – Ore dressing process line as an object with concentrated inputs and distributed outputs

Taking into account the above, the global criterion for optimizing the control of the technological line of beneficiation as a dynamic object with concentrated inputs and distributed outputs is written as follows

$$\begin{cases}Q\left(\bar{p},\bar{\psi},\bar{\beta}^d\right)\to\max;\\ \beta_l\le\beta\left(\bar{p},\bar{\psi},\bar{\beta}^d\right)\le\beta_h.\end{cases} \qquad (21)$$

The structure of a two-level coordinated distributed control system for a technological line for the beneficiation of iron ore raw materials as an object with concentrated inputs and distributed outputs is shown in Fig. 4.

At the upper level of system control (Fig. 4), the formation of coordinating variables is carried out $\bar{p}$, that affect the value of $\Delta\beta_i^d$ transformation of the function of distribution of the useful component by particle size classes of ore material at the *i*-th point, i.e., the load on the *i*-th technological unit of the concentration line.

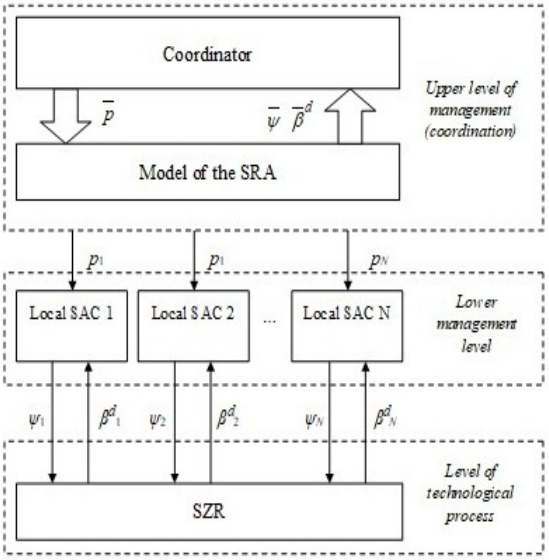



Figure 4 – System of two-level coordinated separated control of the technological line of enrichment

At the lower level of control, concentrated control influences are formed by local automatic control systems for individual technological units and cycles.

## 4 EXPERIMENTS

The control of a distributed system of technological units of an ore dressing line is formed on the basis of decomposition of the dynamics of a distributed system into time and space components [4]. In the spatial domain, the control synthesis problem is solved as a sequence of approximation problems of a set of spatial components of the dynamics of the controlled system. In the time domain, the solution of the control synthesis problem is based on the methods of synthesizing control systems with concentrated parameters. Thus, in the


DOI 10.15588/1607-3274-2024-4-20

structure of the control system, blocks of spatial and temporal formation of control influences are distinguished (Fig. 5).

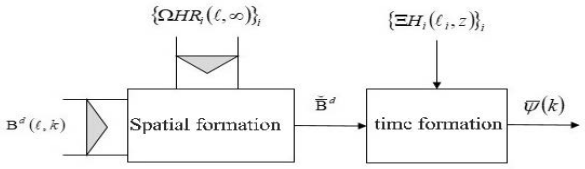



Figure 5 – Structure of the control system for the distributed ore dressing process

To form spatial control, the corresponding component is used $\{\Omega HR_i(\ell,\infty)\,|\,i=\overline{1,n}\}$, obtained as a result of decomposition of the distributed dynamics of the technological enrichment line. At the output of the spatial formation unit, a vector of concentrated variables is formed $\breve{\breve{\mathrm{B}}}^d(\ell)$, which is the basis for the temporal formation of control taking into account the corresponding component $\{\Xi H_i(\ell_i,s)\,|\,i=\overline{1,n}\}$ distributed dynamics of the control object. The output of the time control formation unit is a vector of zoned control variables $\overline{\psi}(k)$ for local control systems of individual technical units.

In this case, the control task is to form the following sequence of control inputs $\overline{\psi}(k)$, which is in a steady state $k\to\infty$, ensure minimal control error at the quadratic norm

$$\|\varepsilon(\ell,\infty)\| = \left\|\mathrm{B}^d(\ell,\infty) - \beta^d(\ell,\infty)\right\| \to \min. \qquad (22)$$

The spatial control formation block (Fig. 5) solves the optimization problem of approximation

$$\left\|\mathrm{B}^d(\ell,\infty) - \sum_{i=1}^{n}\mathrm{B}_i{}^d(\ell_i,\infty)\Omega HR_i(\ell,\infty)\right\| \to \min. \qquad (23)$$

In [11], it was shown that the solution to the approximation problem (23) is guaranteed as a unique best approximation

$$\mathrm{B}^{d*}(\ell,\infty) = \sum_{i=1}^{n}\breve{\mathrm{B}}_i{}^d(\ell_i,\infty)\Omega HR_i(\ell,\infty). \qquad (24)$$

The control vector $\breve{\breve{\mathrm{B}}}^d$ is fed to the input of the time shaping unit, which contains the corresponding number of single-channel control loops with concentrated parameters. Each of the control loops generates the corresponding element of the vector of concentrated control actions $\overline{\psi}(k)$ and consists of a regulator $\{\mathrm{K}_i(z)\,|\,i=\overline{1,n}\}$ and a control object with a zero-order extrapolator $\{\Xi H_i(\ell_i,z)\,|\,i=\overline{1,n}\}$. When forming the control, the individual components of the vector $\breve{\breve{\mathrm{B}}}^d = \{\breve{\mathrm{B}}_i{}^d(\ell_i,\infty)\,|\,i=\overline{1,n}\}$ are inputs of separate control circuits $\{\mathrm{K}_i(z),\Xi H_i(\ell_i,z)\,|\,i=\overline{1,n}\}$.

Thus, the synthesis of control of an open system is carried out in two stages: temporal control – by forming control loops with concentrated parameters; spatial control – by solving the approximation problem.

Time components required for the formation of control of the technological line of ore dressing $\{\Xi H_i(\ell_i,s)\,|\,i=\overline{1,n}\}$ the distributed dynamics of the control object are determined from the matrix of distributed discrete impulse transition functions $\{\Theta H_i(\ell_j,k)\,|\,i=\overline{1,n},\ j=\overline{1,m}\}$. Multiple discrete reviews $\{\Theta H_i(\ell_j,k)\}$ correspond to the set of approximations $\{\Theta H_i'(\ell_j,k)\,|\,i=\overline{1,n},\ j=\overline{1,m}\}$. So, we get

$$\begin{gathered}\beta_i^d(\ell,k) = \Theta H_i'(\ell,k) * \psi_i(k),\ i=\overline{1,n}\\ \left(\beta^d\right)'(\ell,k) = \sum_{i=1}^{n}\Theta H_i'(\ell,k) * \psi_i(k)\\ \left(\beta_i^d\right)'(\ell,k) = \Theta HR_i'(\ell,k)\Theta H_i'(\ell_i,k) * \psi_i(k).\end{gathered} \qquad (25)$$

Since $\beta^d(\ell,k)$ defined as a linear combination of elements of discrete distributional characteristics of the dynamics of the system under consideration $\{\Theta H_i(\ell,k)\}_{i,k}$ – The accuracy of the models depends on the accuracy of the approximation of these characteristics and the physical constraints imposed on a particular control variable $\psi_i^{\max}$. Therefore, to ensure the required accuracy of the model, the differential properties of the set of elements $\{\psi_i^{\max}\Theta H_i(\ell,k)\}_{i,k}$ should be determined using a set of discrete values $\{\Theta H_i(\ell_j,k)\}_{i,j,k}$. These discrete values are used to determine the discrete continuity modulus, as well as the difference norms, which are then generalized over the entire interval $\ell\in[0,L]$ in the corresponding functional spaces. Select the element that has the maximum value of the differential properties and mark it as $\psi_i^{\max}\Theta H^{\max}{}_i(\ell,k)$, setting the approximation accuracy $\chi$ = 5%, and defining the norm || ||, the problem will take the form

$$\left\|\left(\beta^d\right)(\ell,k) - \left(\beta^d\right)'(\ell,k)\right\| \le \chi. \qquad (26)$$

At the same time, it takes into account the limitations imposed on the control variables.


DOI 10.15588/1607-3274-2024-4-20

## 5 RESULTS

The main internal controlling influences in the enrichment line, which is shown in Fig. 1, are water consumption in technological units [4]. Figs. 6–11 show examples of the obtained qualitative and quantitative dependencies that characterize the influence of this parameter on the course of the technological process.

The dependence of the mass fraction of the solid phase in the iron ore pulp on the water flow to the technological units distributed along the concentration line is shown in Fig. 6. The dependence of iron content in the –0.044 mm class and the yield of this class on water consumption is shown in Figs. 7, 8. The dependence of the mass fraction of iron in the by-product on the mass fraction of the solid phase of the pulp distributed along the concentration line is shown in Fig. 9. The dependence of the mass fraction of iron in the industrial product and the iron content in the –0.071 mm class is shown in Fig. 10. 11.

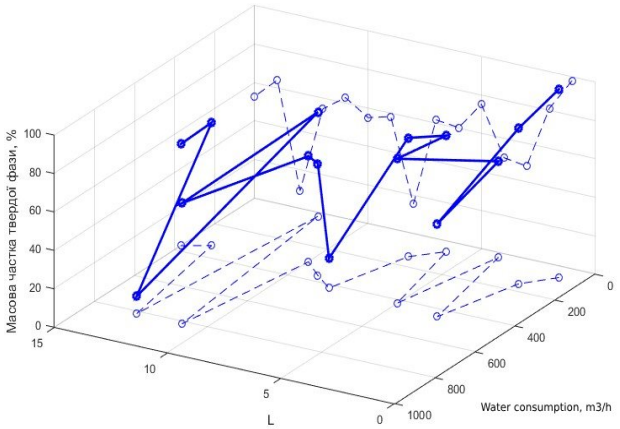



Figure 6 – Mass fraction of the solid phase in iron ore pulp depending on water flow to technological units:
"---" – distributed function; "- - -" – projections on coordinate planes

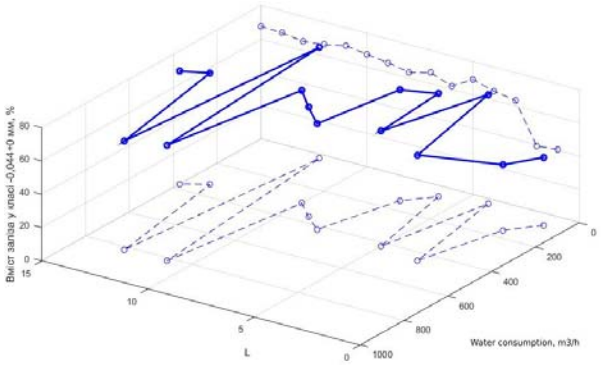



Figure 7 – Iron content in the class "–0.044 mm" depending on the water flow to the technological units:
"---" – distributed function; "- - -" – projections on coordinate planes

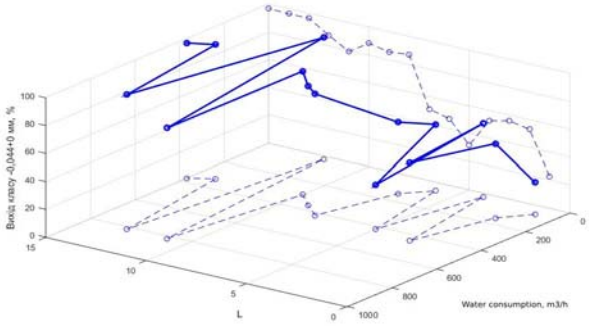



Figure 8 – The output of the size class "–0.044 mm" depending on the water flow to the technological units:
"---" – distributed function; "- - -" – projections on coordinate planes

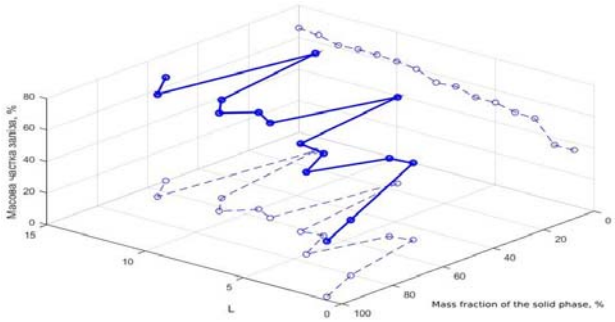



Figure 9 – Mass fraction of iron in the industrial product as a function of the mass fraction of the pulp solid phase:
"---" – distributed function; "- - -" – projections on coordinate planes

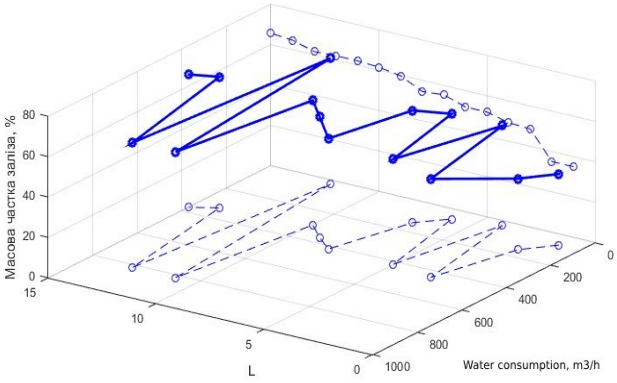



Figure 10 – Mass fraction of iron in industrial products as a function of water flow to technological units:
"---" – distributed function; "- - -" – projections on coordinate planes

## 6 DISCUSSION

Thus, the control of an ore dressing process line as an object with concentrated inputs and distributed outputs involves the synthesis of a spatially distributed control influence, on the basis of which a vector of corresponding time-dependent concentrated influences is formed.


DOI 10.15588/1607-3274-2024-4-20

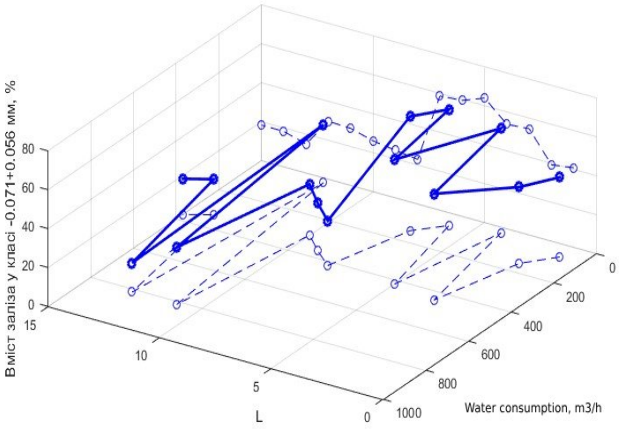



Figure 11 – Iron content in the –0.071 mm class depending on the water flow to the process units:
"---" – distributed function; "- - -" – projections on coordinate planes

The steady-state values of partial distributed transient characteristics are used to calculate the control system. The steady-state values of the transient characteristics are obtained on the basis of the finite value theorem [11, 12].

$$H_i(\ell,\infty)=\lim_{s\to\infty} s\left\{\Xi_i(\ell,s)\frac{1}{s}\right\}=$$
$$\lim_{s\to\infty}\Xi_i(\ell,s)=\lim_{s\to\infty} z\Xi H(\ell,z)=H_i(\ell,\infty),$$
$$H_i(\ell,\infty)=\Xi A_i(0)\Xi G_i(0)\int_0^L \Xi(\ell,\xi,0)T_i(\xi)d\xi. \qquad (27)$$

Then the following feedback was given

$$HR_i(\ell,\infty)=\frac{H_i(\ell,\infty)}{H_i(\ell,\infty)}=\frac{H_i(\ell,\infty)}{H_i(\ell,\infty)}=HR_i(\ell,\infty). \qquad (28)$$

The results of calculating the constant values are shown in Fig. 12.

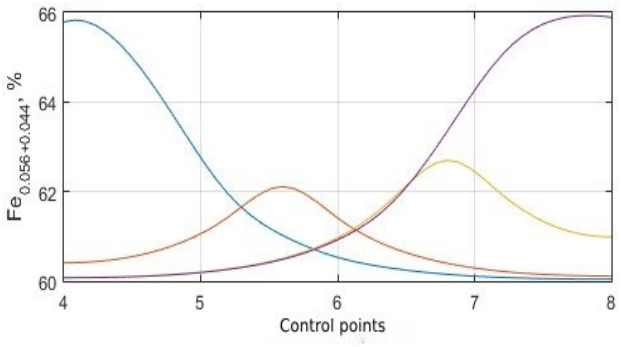



Figure 12 – Steady-state values of transient characteristics for individual inputs according to the control points of Fig. 1.

The given responses are determined as follows: first, the maximum functional values of the responses at the points $\{l_j\}_{j=1,11}$. Then the values at specific points $\{l_j\}_{j=1,11}$ divided by this value. The results of the reduction are shown in Fig. 13.

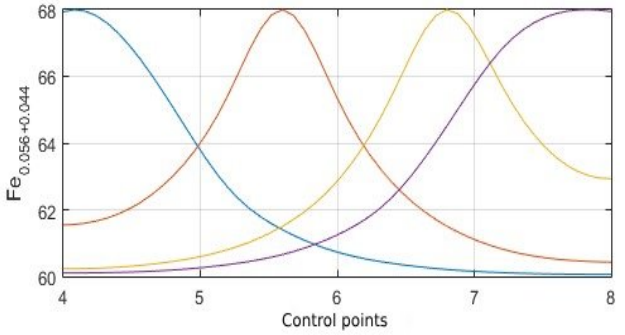



Figure 13 – Steady-state transient characteristics for individual inputs according to
to the control points of Fig. 1.

When forming a sequence of control actions $\psi_i(k)$ robust regulators used $K_i(z)$. The optimal parameters of the controllers were determined in accordance with the results obtained in [4].

The results obtained allow us to conclude that it is expedient to represent the technological line for the beneficiation of iron ore raw materials as a structure with concentrated inputs and distributed outputs and to use appropriate approaches to form the control of the system of technological units of mining production.

The testing of this approach in the formation of automated control of technological enrichment lines at ore dressing plants of the Kryvyi Rih iron ore basin shows that its application allows to reduce by 2–5% the fluctuations in the material composition of enrichment products, that go to the metallurgical processing, increase the content of useful component in the concentrate by 0.3–0.8%, maximize the productivity of technological units in conditions of variaty of static and dynamic characteristics of control objects and reduce by 0.2 – 0.9% of energy consumption.

## CONCLUSIONS

The technological line of enrichment is decomposed into a set of separate subsystems (technological units, enrichment cycles). Under these circumstances, the solution of the global optimization problem is also decomposed into a corresponding set of individual subproblems of optimizing the control of subsystems.

The structure of the system of two-level coordinated distributed control of the technological line of iron ore beneficiation as an object with concentrated inputs and distributed outputs is proposed.

The formation of control of the distributed system of technological units of the ore dressing line is based on the decomposition of the dynamics of the distributed system into time and space components. In the spatial domain, the control synthesis problem is solved as a sequence of approximation problems on the set of spatial components of the modeled system dynamics. In the time domain, the solution of the control synthesis problem is based on the methods of synthesizing control systems with concentrated parameters.


DOI 10.15588/1607-3274-2024-4-20

Using the proposed approach at mining enterprises in the Kryvyi Rih iron ore basin will improve the quality of iron ore concentrate supplied to metallurgical processing, increase the productivity of technological units and reduce energy consumption.

**The scientific novelty.** The paper proposes the structure of a two-level coordinated distributed control system for a technological line for the beneficiation of iron ore raw materials as an object with concentrated inputs and distributed outputs. The formation of the control of a distributed system of technological units of the ore dressing line is based on the decomposition of the dynamics of the distributed system into time and space components. In the spatial domain, the control synthesis problem is solved as a sequences of approximation problems on the set of spatial components of the dynamics of the controlled system. In the time domain, the solution of the control synthesis problem is based on the methods of synthesizing control systems with concentrated parameters.

**The practical significance.** Using the proposed approach at mining enterprises of the Kryvyi Rih iron ore basin will improve the quality of iron ore concentrate supplied to metallurgical processing, increase the productivity of technological units and reduce energy consumption

**Prospects for further research.** Prospects for further research are to study the possibility of using the proposed approach to solve a wide class of practical problems.

## ACKNOWLEDGEMENTS

The work was carried out within the framework of the research topic of the Kryvyi Rih National University “Decentralized optimal control of interdependent mining processes based on a dynamic space-time model” (state registration number 0116U001519).

## REFERENCES

1. Wang X., Gu X., Liu Z., Wang Q., Xu X. and Zheng M. Production Process Optimization of Metal Mines Considering Economic Benefit and Resource Efficiency Using an NSGA-II Model, *Processes*, 2018, Vol. 6(11):228. https://doi.org/10.3390/pr6110228.
2. Chai T., Ding J., Yu G., Wang H. Integrated Optimization for the Automation Systems of Mineral Processing, *IEEE Transactions on Automation Science and Engineering*, 2014, Vol. 11, pp. 965–982. https:// doi 10.1109/TASE.2014.2308576.
3. Yang C. & Ding J. Constrained dynamic multi-objective evolutionary optimization for operational indices of beneficiation process, *Journal of Intelligent Manufacturing*, 2019, Vol. 30, pp. 2701–2713. https://doi.org/10.1007/s10845-017-1319-1.
4. Morkun V., Morkun N., Tron V. Distributed control of ore beneficiation interrelated processes under parametric uncertainty, *Metallurgical and Mining Industry*, 2015, Vol. 7, No. 8, pp. 18–21. https://www.metaljournal.com.ua/assets/Journal/english-edition/MMI_2015_8/004Morkun.pdf.
5. Wang C., Ding J., Cheng R., Liu Ch., Chai T. Data-Driven Surrogate-Assisted Multi-Objective Optimization of Complex Beneficiation Operational Process, *20th IFAC World Congress*, 2017, Vol. 50, pp. 14982–14987. https://doi.org/10.1016/j.ifacol.2017.08.2561.
6. Lusty P.A.J., Gunn A. G. Challenges to global mineral resource security and options for future supply, *Geological Society*. London, Special Publications, 2015, Vol. 393, pp. 265–276. https://doi.org/10.1144/SP393.13
7. Tromans D. Mineral comminution: energy efficiency considerations, *Minerals engineering*, 2008, Vol. 21(8), pp. 613–620. https://doi.org/10.1016/j.mineng.2007.12.003
8. Dong W. Distributed optimal control of multiple systems, *International Journal of Control*, 2010, Vol. 83(10), pp. 2067–2079. https://doi.org/10.1080/00207179.2010.504786
9. Wei Wang, Fangfang Zhang & Chunyan Han Distributed LQR control for discrete-time homogeneous systems, *International Journal of Systems Science*, 2016, Vol. 47:15, pp. 3678–3687. https://doi.org/10.1080/00207721.2015.1112931.
10. Mordukhovich B. S. Optimal Control of Distributed Systems, *In: Variational Analysis and Generalized Differentiation II. Grundlehren der mathematischen II. Springer*. Berlin, Heidelberg, 2006, Vol. 331. https://doi.org/10.1007/3-540-31246-3_
11. Hulko G., Antoniova M., Belavy C., Belansky J., Szuda J., Vegh P. Modeling, control and design of distributed parameter systems: internet version of the monograph, *Slovak university of technology Bratislava*, 2003. Access mode: http://www.dpscontrol.sk/monog_index.html.
12. Hulkó G. Control of Lumped Input and Distributed Output Systems, *Preprints of 5-th IFAC, IMACS, IFIP Symposium on Control of Distributed Parameter Systems*. Perpignan, 1989.
13. Tkáčik T., Tkáčik M., Jadlovská S. and Jadlovská A. Modeling and Analysis of Distributed Control Systems, *Processes*, 2024, Vol. 12(1):5. https://doi.org/10.3390/pr12010005.
14. Ge X., Yang F., Han Q.-L. Distributed networked control systems, *A brief overview. Information Sciences*. 2017, Vol. 380, pp. 117–131. https:// doi.org/ 10.1016/j.ins.2015.07.047.
15. Oumeziane F. A., Ourghanlian Al., Amari S. Analysis of distributed control systems using timed automata with guards and dioid algebra, *In Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. Vienna, Austria, 8–11 September 2020, Vol. 1, pp. 1373–1376. https:// doi.org/ 10.1109/ETFA46521.2020.9211898.
16. Li J., Wang Zh., Sun L., Wang W. Modeling and analysis of network control system based on hierarchical coloured Petri net and Markov chain, *Discrete Dynamics in Nature and Society*, 2021, Vol. 1, 9948855. https://doi.org/10.1155/2021/9948855.
17. Baldellon O., Fabre J.-Ch., Roy M. Modeling distributed real-time systems using adaptive petrinets, *Actes de la 1re journée 3SL*. Saint-Malo, France, 2011, Vol. 10, pp. 7–8. Access mode: https://www.univ-orleans.fr/lifo/evenements/3SL/actes/3sl_04.pdf.
18. Zhang X.-M., Han Q.-L., Ge X., Ding D., Ding L., Yue D., Peng Ch. Networked control systems: A survey of trends and techniques, *IEEE/CAA Journal of Automatica Sinica*, 2020, Vol. 7, pp. 1–17. https:// doi.org/ 10.1109/JAS.2019.1911651
19. Cruz E. M., García Carrillo L. R., Member Structuring Cyber-Physical Systems for Distributed Control with IEC 61499 Standard, *IEEE Latin America Transactions*, 2023, Vol. 21, pp. 251–259. https:// doi.org/ 10.1109/TLA.2023.10015217
20. Cai N., Sabatini R., Dong X.-W., Junaid Khan M., and Yu Y. Decentralized Modeling, Analysis, Control, and Application of Distributed Dynamic Systems, *Journal of Control Science and Engineering*, 2016, 8985017. https://doi.org/10.1155/2016/8985017.
21. Hulewicz A. Distributed control system DCS using a PLC controller, *ITM Web of Conferences Computer Applications in Electrical Engineering (ZkwE’2019)*, 2019, Vol. 28, 01041, P. 2. https://doi.org/10.1051/itmconf/20192801041.
22. Jabeen H. Mastering Distributed Control Systems: A Comprehensive Guide to DCS Architecture, Components, and Applications, https://www.wevolver.com/article/mastering-distributed-control-systems-a-comprehensive-guide-to-dcs-architecture-components-and-applications.


DOI 10.15588/1607-3274-2024-4-20

23. Porkuian O., Morkun V., Morkun N., Serdyuk O. Predictive Control of the Iron Ore Beneficiation Process Based on the Hammerstein Hybrid Model, *Acta Mechanica et Automatica*, 2019, Vol. 13(4), pp. 262–270. https:// doi.org/ 10.2478/ama-2019-0036.
24. Bi L., Wang Z., Wu Z. and Zhang Y. A New Reform of Mining Production and Management Modes under Industry 4.0: Cloud Mining Mode, *Applied Sciences*, 2022, Vol. 12, 2781. https://doi.org/10.3390/app12062781
25. Chunbao S., Kuangdi, X., In: Xu K. (eds) Ore Crushing, *The ECPH Encyclopedia of Mining and Metallurgy. Springer, Singapore*, 2023. https://doi.org/10.1007/978-981-19-0740-1_1119-1
26. Wang X. L., Liu L. M., Duan L. and Liao Q. Multi-Objective Optimization for an Industrial Grinding and Classification Process Based on PBM and RSM, *IEEE/CAA Journal of Automatica Sinica*. November, 2023, Vol. 10, No. 11, pp. 2124–2135. https:// doi.org/ 10.1109/JAS.2023.123333.
27. Morkun V., Morkun N., Tron V., Hryshchenko S., Serdyuk O., Dotsenko I. Basic regularities of assessing ore pulp parameters in gravity settling of solid phase particles based on ultrasonic measurements, *Archives of Acoustics*, 2019, Vol. 44(1), pp. 161–167. https:// doi.org/ 10.24425/aoa.2019.126362
28. Morkun V., Morkun N., Pikilnyak A. Ultrasonic facilities complex for grinding and ore classification process control, *2017 IEEE 37th International Conference on Electronics and Nanotechnology, ELNANO 2017 – Proceedings*, 2017, pp. 409–413, 7939788. https:// doi.org/ 10.1109/ELNANO.2017.7939788
29. Young A., Rogers P. A Review of Digital Transformation in Mining, *Mining, Metallurgy & Exploration*, 2019, Vol. 36, pp. 683–699. https:// doi.org/ 10.1007/s42461-019-00103-w
30. Azaryan A. A., Batareyev, A. S., Karamanits F. I., Kolosov B. A., Morkun V. S. Ways to reduce ore losses and dilution in iron ore underground mining in Kryvbass, *Science and Innovation*, 2018, Vol. 14(4), pp. 17–24. https:// doi.org/ 10.15407/scine14.03.017.
31. Xiaoping L., Dayong Z., Kuangdi X. Measurement and Control for Ore Grinding and Classification, *The ECPH Encyclopedia of Mining and Metallurgy*. Springer, Singapore, 2023. https://doi.org/10.1007/978-981-19-0740-1_830-1.
32. Morkun V. S., Morkun N. V., Tron V. V. Automatic control of the ore suspension solid phase parameters using high-energy ultrasound, *Radio electronics computer science control*, 2017, No. 3, pp. 175–182. https://doi.org/10.15588/1607-3274-2017-3-19.

Received 30.08.2024.
Accepted 31.10.2024.

УДК 622.7: 534

## ЗАГАЛЬНІ ПРИНЦИПИ ФОРМАЛІЗАЦІЇ КЕРУВАННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ ГІРНИЧОГО ВИРОБНИЦТВА ЯК ДИНАМІЧНОЮ РОЗПОДІЛЕНОЮ СИСТЕМОЮ

**Моркун В. С.** – д-р техн. наук, проф., професор Байротського університету, Байрот, Німеччина.
**Моркун Н. В. –** д-р техн. наук, проф., професор Байротського університету, Байрот, Німеччина.
**Грищенко С. М. –** канд. пед. наук, старший дослідник, доцент кафедри комп'ютерних та інформаційних технологій і систем Державного податкового університету, Ірпінь, Україна.
**Шашкіна А. А. –** аспірант Криворізького національного університету, Кривий Ріг, Україна.
**Бобров Є. Ю. –** аспірант Криворізького національного університету**,** Кривий Ріг, Україна.

### АНОТАЦІЯ

**Актуальність.** Проблема синтезу, моделювання та аналізу автоматизованого управління складними технологічними процесами гірничого виробництва як динамічної структури з розподіленими параметрами.

**Мета роботи.** На прикладі технологічної лінії збагачення руди розглянути загальні принципи формалізації управління процесами гірничого виробництва як динамічної системи з розподіленими параметрами.

**Метод**. Моделювання взаємодій між окремими компонентами системи керування здійснено з використанням методів скоординованого розподіленого керування. Відповідно до цього підходу технологічна лінія декомпозується на сукупність окремих підсистем (технологічних агрегатів, циклів збагачення). За таких умов розв'язання глобальної оптимізаційної задачі також декомпозується на відповідну множину окремих підзадач оптимізації управління підсистемами. Для вирішення глобальної задачі в цій постановці використовується дворівнева структура з координуючими змінними, які подаються на вхід локальних систем управління технологічними агрегатами і циклами. На нижньому рівні управління множини підзадач мають незалежні розв'язки, які координуються координуючими змінними, сформованими на верхньому рівні.

**Результати.** У статті запропоновано метод формування управління розподіленою системою технологічних агрегатів гірничо-збагачувальної лінії на основі декомпозиції динаміки розподіленої системи на часову та просторову складові. У просторовій області задача синтезу керування вирішується як послідовність задач апроксимації множини просторових складових динаміки керованої системи. У часовій області розв'язання задачі синтезу керування базується на методах синтезу систем керування із зосередженими параметрами.

**Висновки.** Використання запропонованого підходу до формування системи управління технологічним процесом на гірничодобувних підприємствах Криворізького залізорудного басейну дозволить підвищити якість залізорудного концентрату, що надходить на металургійну переробку, збільшити продуктивність технологічних агрегатів і знизити енергоспоживання.

**КЛЮЧОВІ СЛОВА:** видобуток, автоматизація, збагачення руди, розподілене управління, процес, система.

### ЛІТЕРАТУРА

1. Wang X. Production Process Optimization of Metal Mines Considering Economic Benefit and Resource Efficiency Using an NSGA-II Model / [Xunhong Wang, Xiaowei Gu, Zaobao Liu, Qing Wang et al.] // Processes. – 2018. – Vol. 6(11):228. https://doi.org/10.3390/pr6110228.
2. Integrated Optimization for the Automation Systems of Mineral Processing / [T. Chai, J. Ding, G. Yu, H. Wang] // IEEE Transactions on Automation Science and Engineering. – 2014. – Vol. 11. – P. 965–982. https:// doi 10.1109/TASE.2014.2308576.
3. Yang C. Constrained dynamic multi-objective evolutionary optimization for operational indices of beneficiation process / Cuie Yang & Jinliang Ding // Journal of Intelligent Manufacturing. – 2019. – Vol. 30. – P. 2701–2713. https://doi.org/10.1007/s10845-017-1319-1.
4. Morkun V. Distributed control of ore beneficiation interrelated processes under parametric uncertainty / Vladimir Morkun, Natalia Morkun, Vitaliy Tron // Metallurgical and Mining Industry. – 2015. – Vol. 7, No. 8. – P. 18–21.


DOI 10.15588/1607-3274-2024-4-20

https://www.metaljournal.com.ua/assets/Journal/english-edition/MMI_2015_8/004Morkun.pdf.
5. Wang C. Data-Driven Surrogate-Assisted Multi-Objective Optimization of Complex Beneficiation Operational Process / [Chengzhi Wang, Jinliang Ding, Ran Cheng et al.]// 20th IFAC World Congress. – 2017. – Vol. 50. – P. 14982–14987. https://doi.org/10.1016/j.ifacol.2017.08.2561.
6. Lusty P.A.J. Challenges to global mineral resource security and options for future supply / P. A. J. Lusty, A. G. Gunn // Geological Society, London, Special Publications. – 2015. – Vol. 393. – P. 265–276. https://doi.org/10.1144/SP393.13
7. Tromans D. Mineral comminution: energy efficiency considerations / Desmond Tromans // Minerals engineering. – 2008. – Vol. 21(8). – P. 613–620. https://doi.org/10.1016/j.mineng.2007.12.003
8. Dong W. Distributed optimal control of multiple systems / Wenjie Dong // International Journal of Control. – 2010. – Vol. 83(10). – P. 2067–2079. https://doi.org/10.1080/00207179.2010.504786
9. Wei Wang Distributed LQR control for discrete-time homogeneous systems / Wei Wang, Fangfang Zhang & Chunyan Han // International Journal of Systems Science. – 2016. – Vol. 47:15. – P. 3678–3687. https://doi.org/10.1080/00207721.2015.1112931.
10. Mordukhovich B. S. Optimal Control of Distributed Systems / Boris S. Mordukhovich // In: Variational Analysis and Generalized Differentiation II. Grundlehren der mathematischen II. – Springer, Berlin : Heidelberg. – 2006. – Vol. 331. https://doi.org/10.1007/3-540-31246-3_
11. Hulko G. Modeling, control and design of distributed parameter systems: internet version of the monograph / [G. Hulko, M. Antoniova, C. Belavy et al.] // Slovak university of technology Bratislava. – 2003. Access mode: http://www.dpscontrol.sk/monog_index.html.
12. Hulkó G. Control of Lumped Input and Distributed Output Systems. / G. Hulkó // Preprints of 5-th IFAC, IMACS, IFIP Symposium on Control of Distributed Parameter Systems. – Perpignan. – 1989.
13. Modeling and Analysis of Distributed Control Systems / [Tomáš Tkáčik, Milan Tkáčik, Slávka Jadlovská and Anna Jadlovská] // Processes. – 2024. – Vol. 12(1):5. https://doi.org/10.3390/pr12010005.
14. Ge X. Distributed networked control systems / Xiaohua Ge, Fuwen Yang, Qing-Long Han // A brief overview. Information Sciences. – 2017. – Vol. 380. – P. 117–131. https:// doi.org/ 10.1016/j.ins.2015.07.047.
15. Oumeziane F. A. Analysis of distributed control systems using timed automata with guards and dioid algebra / Fatima Ait Oumeziane, Alain Ourghanlian, Saïd Amari // In Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). – Vienna, Austria, 8–11 September 2020. – Vol. 1. – P. 1373–1376. https:// doi.org/ 10.1109/ETFA46521.2020.9211898.
16. Modeling and analysis of network control system based on hierarchical coloured Petri net and Markov chain / [Jingdong Li, Zhangang Wang, Liankun Sun, Wanru Wang] // Discrete Dynamics in Nature and Society. – 2021. – Vol. 1. – 9948855. https://doi.org/10.1155/2021/9948855.
17. Baldellon O. Modeling distributed real-time systems using adaptive petri nets / Olivier Baldellon, Jean-Charles Fabre, Matthieu Roy // Actes de la 1re journée 3SL. – Saint-Malo, France. – 2011. – Vol. 10. – P. 7–8. Access mode: https://www.univ-orleans.fr/lifo/evenements/3SL/actes/3sl_04.pdf.
18. Zhang X.-M. Networked control systems: A survey of trends and techniques / [Xian-Ming Zhang, Qing-Long Han, Xiaohua Ge et al.] // IEEE/CAA Journal of Automatica Sinica. – 2020. – Vol. 7. – P. 1–17. https:// doi.org/ 10.1109/JAS.2019.1911651
19. Cruz E. M. Structuring Cyber-Physical Systems for Distributed Control with IEC 61499 Standard / E. Monroy Cruz, L. R. García Carrillo, Member // IEEE Latin America Transactions. – 2023, – Vol. 21. – P. 251–259. https:// doi.org/ 10.1109/TLA.2023.10015217
20. Cai N. Decentralized Modeling, Analysis, Control, and Application of Distributed Dynamic Systems / [Ning Cai, Roberto Sabatini, Xi-Wang Dong et al.] // Journal of Control Science and Engineering. – 2016. – 8985017. https://doi.org/10.1155/2016/8985017.
21. Hulewicz A. Distributed control system DCS using a PLC controller / A. Hulewicz // ITM Web of Conferences Computer Applications in Electrical Engineering (ZkwE'2019). – 2019. – Vol. 28. 01041. – P. 2. https://doi.org/10.1051/itmconf/20192801041.
22. Jabeen H. Mastering Distributed Control Systems: A Comprehensive Guide to DCS Architecture, Components, and Applications / Hafsa Jabeen: https://www.wevolver.com/article/mastering-distributed-control-systems-a-comprehensive-guide-to-dcs-architecture-components-and-applications.
23. Predictive Control of the Iron Ore Beneficiation Process Based on the Hammerstein Hybrid Model / [Olga Porkuian, Vladimir Morkun, Natalia Morkun, Oleksandra Serdyuk] // Acta Mechanica et Automatica. – 2019. – Vol. 13(4). – P. 262–270. https:// doi.org/ 10.2478/ama-2019-0036.
24. Bi L. A New Reform of Mining Production and Management Modes under Industry 4.0: Cloud Mining Mode / [Lin Bi, Zhuo Wang, Zhaohao Wu and Yuhao Zhang] // Applied Sciences. – 2022. – Vol. 12. – 2781. https://doi.org/10.3390/app12062781
25. Chunbao S. Ore Crushing / Chunbao S., Kuangdi, X. // In: Xu, K. (eds). The ECPH Encyclopedia of Mining and Metallurgy.– Springer, Singapore. – 2023. https://doi.org/10.1007/978-981-19-0740-1_1119-1
26. Multi-Objective Optimization for an Industrial Grinding and Classification Process Based on PBM and RSM / [X. L. Wang, L. M. Liu, L. Duan and Q. Liao] // in IEEE/CAA Journal of Automatica Sinica. November. – 2023. – Vol. 10, No. 11. – P. 2124–2135. https:// doi.org/ 10.1109/JAS.2023.123333.
27. Basic regularities of assessing ore pulp parameters in gravity settling of solid phase particles based on ultrasonic measurements / [V. Morkun, N. Morkun, V. Tron, S. Hryshchenko et al.]// Archives of Acoustics. – 2019. – Vol. 44(1). – P. 161–167. https:// doi.org/ 10.24425/aoa.2019.126362
28. Morkun V. Ultrasonic facilities complex for grinding and ore classification process control / Vladimir Morkun, Natalia Morkun, Andrey Pikilnyak // 2017 IEEE 37th International Conference on Electronics and Nanotechnology, ELNANO 2017 – Proceedings. – 2017. – P. 409–413, 7939788. https:// doi.org/ 10.1109/ELNANO.2017.7939788
29. Young A. A Review of Digital Transformation in Mining / A. Young, P. Rogers // Mining, Metallurgy & Exploration. – 2019. – Vol. 36. – P. 683–699. https:// doi.org/ 10.1007/s42461-019-00103-w
30. Ways to reduce ore losses and dilution in iron ore underground mining in Kryvbass / [A. A. Azaryan, A. S. Batareyev, F. I. Karamanits et al.] // Science and Innovation, – 2018. – Vol. 14(4). – P. 17–24. https:// doi.org/ 10.15407/scine14.03.017.
31. Xiaoping L. Measurement and Control for Ore Grinding and Classification / L. Xiaoping, Z. Dayong, X. Kuangdi // The ECPH Encyclopedia of Mining and Metallurgy. Springer, Singapore. – 2023. https://doi.org/10.1007/978-981-19-0740-1_830-1.
32. Morkun V. S. Automatic control of the ore suspension solid phase parameters using high-energy ultrasound / V. S. Morkun, N. V. Morkun, V. V. Tron // Radio electronics computer science control. – 2017. – No. 3, — P. 175–182. https://doi.org/10.15588/1607-3274-2017-3-19.


DOI 10.15588/1607-3274-2024-4-20

*Наукове видання*

# Радіоелектроніка,

# інформатика,

# управління

№ 4/2024

Науковий журнал

Головний редактор – д-р техн. наук С. О. Субботін

Заст. головного редактора – д-р техн. наук Д. М. Піза

Комп’ютерне моделювання та верстання — С. В. Зуб
Редактор англійських текстів — С. О. Субботін

Оригінал-макет підготовлено у редакційно-видавничому відділі НУ «Запорізька політехніка»

Реєстрація суб’єкта у сфері друкованих медіа:
Рішення Національної ради України з питань телебачення і радіомовлення № 3040 від 07.11.2024 року
Ідентифікатор медіа: R30-05582

*Підписано до друку 21.11.2024. Формат 60×84/8.*
*Папір офс. Різогр. друк. Ум. друк. арк. 25,8.*
*Тираж 300 прим. Зам. № 1323.*

*69063, м. Запоріжжя, НУ «Запорізька політехніка», друкарня, вул. Жуковського, 64*

Свідоцтво суб’єкта видавничої справи
ДК № 6952 від 22.10.2019.