МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Одеська політехніка»

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Том 14, № 4

Volume 14, No. 4

Одеса – 2024
Odesa – 2024

Журнал внесений до переліку наукових фахових видань України (технічні науки) згідно наказу Міністерства освіти і науки України № 463 від 25.04.2013 р. Перереєстровано на категорію «Б» за фахами 121, 122, 125, 151 згідно наказу МОН України № 1473 від 26.11.2020 р.

Виходить 4 рази на рік

Заснований Одеським національним політехнічним університетом у 2011 році

Свідоцтво про державну реєстрацію КВ № 17610 - 6460Р від 04.04.2011р.

Published 4 times a year

Founded by Odesa National Polytechnic University in 2011

Certificate of State Registration КВ № 17610 - 6460Р of 04.04.2011

Друкується за рішенням редакційної колегії та Вченої ради Національного університету «Одеська політехніка»

Оригінал-макет виготовлено редакцією журналу

Адреса редакції: 1, Шевченка пр., Одеса, 65044, Україна
Телефон: +38 048 705 8506
Web: www.immm.op.edu.ua (immm.opu.ua) Email: immm.ukraine@gmail.com

Editorial address: 1, Shevchenko Ave., Odesa, 65044, Ukraine
Tel.: +38 048 705 8506
Web: www.immm.op.edu.ua (immm.opu.ua) Email: immm.ukraine@gmail.com

ЗМІСТ/CONTENTS

DOI 10.15276/imms.v14.no4.273
UDC 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 273-283

STEGANALYSIS OF A METHOD WITH CODE CONTROL OF INFORMATION EMBEDDING IN THE WALSH-HADAMARD TRANSFORM DOMAIN

O. O Lanovska[1], A. V. Sokolov[2]

[1]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044 Ukraine
[2]National University Odesa Law Academy
23, Fontanska road, Odesa, 65009, Ukraine
Email: radiosquid@gmail.com

Steganography is an integral part of modern information protection systems. The steganographic method with code control of information embedding is a modern steganographic method that operates in the spatial domain of the container. The advantages of this method include: ensuring the reliability of perception, resistance to attacks against the embedded message, sufficient bandwidth, and high computational efficiency. In contrast to steganographic methods, steganalysis methods are created to allow the detection of embedded information. To date, no research has been performed on the resistance of the steganographic method with code control to cryptanalysis attacks. The specific methods for detecting interference using the steganographic method with code control are also unknown. The purpose of this paper is to develop a steganalysis method for detecting a covert channel organized using the steganographic method with code control of information embedding in the Walsh-Hadamard transform domain. In the paper, the research on the behavior of the Walsh-Hadamard transformants of containers and steganographic messages is performed, which allowed us to formulate the conditions for the presence of additional information in the image. These conditions became the basis for the development of an efficient and mathematically simple steganographic method that uses the Walsh-Hadamard transform domain. The performed research on the proposed method allowed us to establish its high efficiency in various conditions with low computational complexity. In particular, it is shown that the efficiency of the proposed method exceeds the efficiency of the histogram analysis method and methods implemented in the well-known StegExpose tool. The results obtained allow us to recommend the proposed steganographic method for practical application for detecting covert communication channels created using the steganographic method with code control. In particular, the simplicity of its algorithmic implementation makes the proposed method effective in conditions of constrained computing resources.

Keywords: steganography, Walsh-Hadamard transform, steganalysis, code control of information embedding.

Introduction and statement of the problem. The constant development of information technologies and their integration into all areas of societal activities occurs in the modern World. This increases the importance of information security systems, while the growing share of multimedia information in global traffic leads to the rising significance of steganographic methods in information protection systems. These methods can conceal the very fact of the protected information's existence. New steganographic methods are continually evolving and improving, which, in turn, emphasizes the importance of developing steganalysis methods in parallel. In situations where the transmission of information through covert channels can become critically important, the development of fast and computationally efficient methods for detecting these covert channels becomes crucial. These methods should consider the specifics of particular techniques to achieve more accurate results.

At the moment, there are many steganographic methods, but they are mainly divided into two categories: the first category involves hiding data in the spatial domain, and the second — in the domain of container transformants.

Methods operating in the spatial domain of the container, for example, the classical LSB method, are characterized by high computational efficiency, high bandwidth, and the ability to easily ensure the reliability of perception, which, despite their simplicity, makes them quite widespread in practice. As the modern versions of the implementation of the LSB method we can consider, for example, the method [1], the main focus of which is on ensuring high bandwidth of the covert channel; method [2], depending on the marking algorithm of connected components; a method [3] that hides more of the secret message in the (sharpest) edges of the image, etc. Such a modification of the LSB methods as LSB-matching (LSBM) also must be considered [4]. The disadvantages of most methods working in the spatial domain of the container include their instability to attacks against the embedded message (for example, compression or noise attacks), and instability to steganalysis.

The methods applying container transform domains can be based on discrete cosine transform (DCT) [5, 6], discrete wavelet transform (DWT) [7], integer wavelet transform (IWT) [8], discrete Fourier transform (DFT) [9].

A characteristic feature of these methods is the preliminary transformation of the container or its blocks into a selected transformation domain, followed by the execution of embedding of additional information. A notable aspect of most of these methods is their ability to provide resilience against attacks targeting the embedded message. However, the use of transformation domains significantly reduces the computational efficiency of these methods, which greatly complicates their implementation on resource-constrained platforms.

The steganographic method with code control of information embedding, which is proposed in [10] is a recent achievement in steganography that has proved its effectiveness. This method is characterized by performing steganographic transformation in the spatial domain of the container, with the capability of selectively influencing the required frequency components of the container blocks. This approach combines the advantages of methods operating in the spatial domain with those methods that operate in the transformation domains: providing resistance to attacks on the embedded message, significant computational efficiency, and ensuring the reliability of perception.

Despite the high prospects and practical application, today the resistance of the steganographic method with code control to steganalysis attacks remains poorly researched, specific methods for detecting covert channels of information transmission, organized using a steganographic method with code control of additional information embedding, are unknown.

As the research performed in this paper shows, the steganographic method with code control of information embedding remains resistant to known steganalysis tools, however, the application of the properties of the Walsh-Hadamard transform opens up prospects for the development of a mathematically simple method for detecting a covert channel based on the steganographic method with code control of information embedding.

The purpose of this paper is to develop a steganalysis method for detecting a covert channel organized using the steganographic method with code control of information embedding in the Walsh-Hadamard transform domain.

This paper is organized as follows: Section 2 provides an overview of the steganographic method with code control. Section 3 explains the proposed steganalysis method. Section 4 presents a comparison of results with other existing methods, while conclusions and suggestions for further research are presented in Section 5.

Steganographic method with code control of information embedding. The foundation of the steganographic method with code control of information embedding lies in the correspondence between the discrete cosine transform (DCT) and the Walsh-Hadamard transform. The core idea of this method is based on utilizing the linearity property of the Walsh-Hadamard transform [11], which enables the embedding of additional information in the spatial domain of the container while targeting a specified frequency component. This method architecture ensures significant computational efficiency, high resistance to attacks against the embedded message, ensuring the reliability of perception, and adequate bandwidth.

Let's introduce the key definitions necessary for our research. The DCT is defined by the following relation

$$S = C_N X C_N^T, \tag{1}$$

where X is the $N \times N$ block of the original image,

C_N^T is the $N \times N$ DCT matrix, the elements $C(i, j)$ of which are calculated using the following equation

$$C(i,j) = \begin{cases} \dfrac{1}{\sqrt{N}}, \text{ when } i = 0; \\ \sqrt{\dfrac{2}{N}} \cos(2j+1)i\pi, \text{ when } i > 0. \end{cases} \tag{2}$$

Another promising type of discrete transform used in the tasks of steganography and steganalysis is the discrete Walsh-Hadamard transform. In matrix form, the one-dimensional version of the Walsh-Hadamard transform can be written as the following matrix product

$$V = YH_N, \tag{3}$$

where Y is the line-vector of length N,

H_N is a Walsh-Hadamard matrix of order $N = 2^k$, which can be constructed following Sylvester's construction, which is represented by the following equation

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}. \tag{4}$$

At that time, the two-dimensional discrete Walsh-Hadamard transform is defined as follows

$$W = H'_N X H'^T_N, \tag{5}$$

where $H'_N = \dfrac{1}{\sqrt{N}} H_N$ is the normalized Walsh-Hadamard matrix,

X is the matrix of the $N \times N$ size.

As the elements of the Walsh-Hadamard matrix are the numbers from $\{-1,1\}$, the Walsh-Hadamard transform is computationally more efficient than the DFT, DCT, and DST transforms [12].

The Walsh-Hadamard transform returns a sequence of values, which represents a generalized notion of frequency.

In [13], a relationship was established between the two-dimensional and one-dimensional Walsh-Hadamard transform. According to this relationship, the coefficients of the two-dimensional Walsh-Hadamard transform can be determined, up to a scaling factor $\dfrac{1}{N}$, using the one-dimensional Walsh-Hadamard transform.

$$W = XH_{N^2}, \tag{6}$$

where the operation A represents a vector of length N^2 obtained by sequentially concatenating the rows of the matrix A of size $N \times N$.

In the research [10], a relationship was identified between the transform matrix of the Walsh-Hadamard transform, the DCT transform, and the components of the singular value decomposition (SVD) of the original matrix. These results provided the theoretical foundation for developing the method with code control of information embedding.

The essence of code control of information embedding lies in ensuring the desired properties of the steganographic message in the spatial domain with minimal computational costs and disturbances introduced to the container during additive embedding of ± 1.

In this approach, one bit of additional information is embedded into each container block, distributed uniformly among the elements of the block.

Let the block $X = \|x_{i,j}\|, i, j, = 0,1,..., N-1$ of a digital image be a matrix of size $N \times N$ while d is the additional information bit needs to be embedded into this image block. A codeword T of size $N \times N$ is assigned to this bit, and used to embed the bit d.

Then, the steganographic message block M will have the form

$$\tilde{M} = \tilde{X} + \tilde{T}. \tag{7}$$

Considering the connection between one-dimensional and two-dimensional Walsh-Hadamard transforms the Walsh-Hadamard transformants of the resulting vector M

$$\tilde{W} = \tilde{M} H_{N^2} = (\tilde{X} + \tilde{T}) H_{N^2} = \tilde{X} H_{N^2} + \tilde{T} H_{N^2}. \tag{8}$$

Expression (8) allows for a fundamental conclusion about the nature of perturbations in the Walsh-Hadamard transform coefficients within the steganographic message after the additive embedding of additional information.

The magnitude and localization of such perturbations depend on the specific form of the term $\tilde{T} H_{N^2}$, which represents the Walsh-Hadamard transformants of the row vector $\tilde{T}$ used to encode the additional information bit d.

Therefore, to implement code control of information embedding its bits must first be encoded with codewords of size $N \times N$ that enable selective influence on specific Walsh-Hadamard transform coefficients and, consequently, on the DCT transformants.

Such pre-coding allows for focused influence on a given Walsh-Hadamard transformant of a selected block of size $N \times N$ while limiting the impact on each container element to a unit amplitude.

This ensures the desired properties of the steganographic transformation depending on which transformant the selected codeword is targeted.

As codewords that provide selective influence on specific Walsh-Hadamard transformant, the matrix representation of the rows of the Walsh-Hadamard matrix of order N^2 is used. As mentioned above, based on the connection between the DCT, it is most appropriate to use codewords that affect the low-frequency transformants. For the DCT, these are the transformants (2,1), (2,2), (1,2), and the transformant (1,1).

Therefore, for the Walsh-Hadamard transform [11], these will correspond to the transformants (5,1), (5,5), (1,5), and the transformant (1,1), respectively. Table 1 presents the most commonly used codewords of order 8×8 that influence the low-frequency and mid-frequency components of the container blocks, as well as their corresponding Walsh-Hadamard transform matrices, up to a scaling factor $1/N$.

Thus, it can be seen how steganographic method with code control influences the Walsh-Hadamard transformants while operating in the spatial domain. This leads to the natural conclusion regarding the feasibility of using the Walsh-Hadamard transform domain for the steganalysis of this method, as the changes occurring in this domain are the most specific and noticeable.

Proposed steganalysis method. For the computational experiment, the main dataset of 530 digital images from the NRCS database in lossless TIFF format was used. An additional dataset of 470 images in lossy JPEG format was also included.

For each dataset, steganographic message sets were created using the codewords affecting (5,1), (5,5), and (1,5) transformants. Additional information was embedded in the red channel.

To determine the criteria that could be applied for detecting interference, the values of the Walsh-Hadamard transformants for the steganographic messages and original containers were analyzed, as obtained according to (7).

The performed research led to the conclusion that steganographic transformation does not cause the amplitude values of the Walsh-Hadamard transformants to exceed the limits typical for the original images. However, for blocks of size 8×8 the following pattern was found, which we will write in the form of the following statement.

Table 1.

Mapping of codewords and their Walsh-Hadamard transformant matrices

Codeword	Walsh-Hadamard transformants
$T^{+}_{8(1,1)}=\begin{bmatrix}1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\end{bmatrix}$	$W_{(1,1)}=\begin{bmatrix}64&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$
$T^{+}_{8(5,1)}=\begin{bmatrix}1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1\\-1&-1&-1&-1&-1&-1&-1&-1\\-1&-1&-1&-1&-1&-1&-1&-1\\-1&-1&-1&-1&-1&-1&-1&-1\\-1&-1&-1&-1&-1&-1&-1&-1\end{bmatrix}$	$W_{(5,1)}=\begin{bmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\64&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$
$T^{+}_{8(1,5)}=\begin{bmatrix}1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\end{bmatrix}$	$W_{(1,5)}=\begin{bmatrix}0&0&0&0&64&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$
$T^{+}_{8(5,5)}=\begin{bmatrix}1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\1&1&1&1&-1&-1&-1&-1\\-1&-1&-1&-1&1&1&1&1\\-1&-1&-1&-1&1&1&1&1\\-1&-1&-1&-1&1&1&1&1\\-1&-1&-1&-1&1&1&1&1\end{bmatrix}$	$W_{(5,5)}=\begin{bmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&64&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$

Statement 1. In the blocks of a steganographic message of size 8×8, the Walsh-Hadamard transformant, where the additional information was embedded, is more likely to take the maximum value in the block compared to the original images. This dependence can be represented by the following formula

$$U_{orig}(k,l,c) < U_{steg}(k,l,c), \tag{9}$$

where c is the color channel, and U_{orig} and U_{steg} are defined by the following relation

$$\begin{gathered} U_{orig}(k,l,c) = P(W_{orig}(k,l,c)) = \max_{i,j}\{W_{orig}(i,j,c)\}), \\ U_{steg}(k,l,c) = P(W_{steg}(k,l,c) = \max_{i,j}\{W_{steg}(i,j,c)\}), \quad (i,j,c) \neq (1,1,c). \end{gathered}, \tag{10}$$

where W_{orig} and W_{steg} are the matrices of the Walsh-Hadamard transformants for the original and steganographic images, respectively.

$P(W_{orig}(k,l,c)) = \max_{i,j}\{W_{orig}(i,j,c)\})$ is the probability of the event that the Walsh-Hadamard transformant with index (i,j) of the block in the original image acquires the maximum value;

$P(W_{steg}(k,l,c) = \max_{i,j}\{W_{steg}(i,j,c)\})$ is the probability of the event that the transformant of the Walsh-Hadamard transform with the index (i,j) of the steganographic message block acquires the maximum value.

The transformant (1,1) is not taken into account, as it is zero-frequency and always takes a value significantly higher than the other transformants.

At the same time, the frequency of maximum values occurring in all other Walsh-Hadamard transformants of the block, except for the one influenced by the codeword, decreases in the steganographic message.

To research the values of U_{orig} and U_{steg}, experiments were performed, and the results are shown in Fig 1 and Fig.2. Using the results, which demonstrate certain patterns of behavior in images during the embedding of information, it is possible to define specific criteria by which an image can be identified as one in which information has been embedded using a steganographic method with code control. Additionally, it is possible to not only detect the presence of interference but also determine the specific channel of embedding and the codeword used with the help of the detected index that disrupts the patterns of the original containers.

So, first of all, the average threshold values of the frequency of occurrence of maximum were determined for each index of the 8×8 block and each channel. However, these values are averaged, which means that it is important to adjust them experimentally. For this, it was decided to focus on the red color channel, namely on the Walsh-Hadamard transformants with indexes (5,1), (5,5), (1,5), in which it is recommended to embed information.

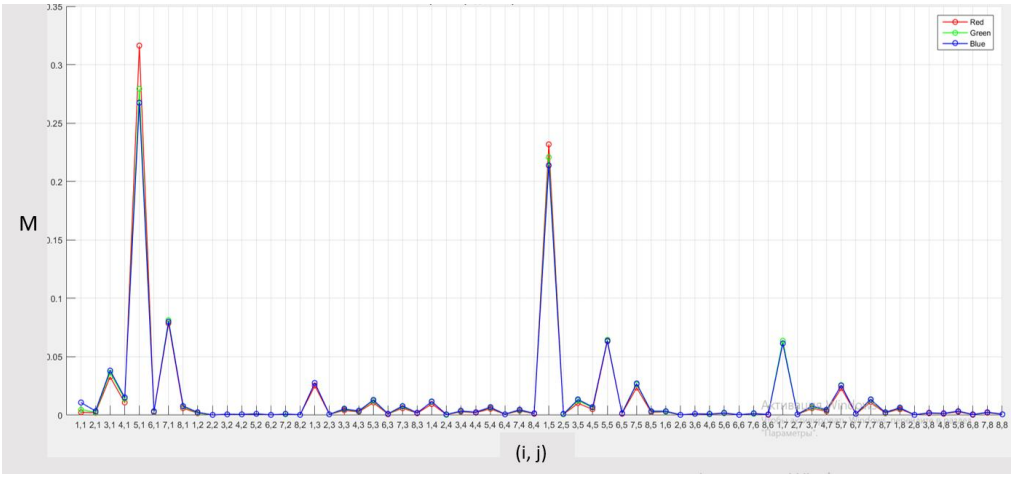


Fig. 1. Graph of the frequency of the location of the maximum value by the index for images without embedded information

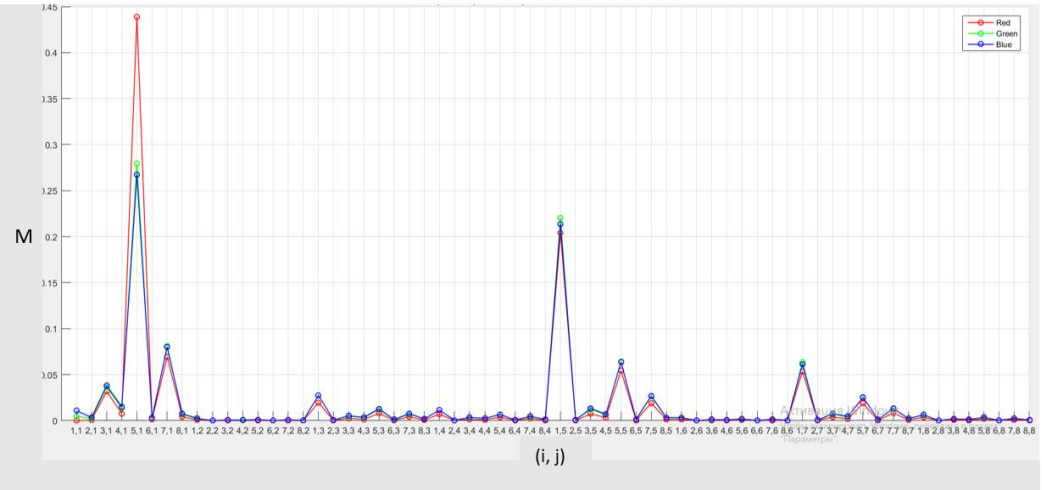


Fig. 2. Graph of the frequency of the location of the maximum value by the index for images with embedded information

To determine the initial matrix of threshold values for each color channel of the set of original images, the matrix of the frequency of occurrence of maximums by index was found. These matrices are formed according to the following formula

$$M(i,j,c)=\frac{1}{N}\sum_{k=1}^{N}\delta\left((i,j,c),\max_{i_{\max},j_{\max},c}W_k(i_{\max},j_{\max},c)\right),$$
$$(i,j)\neq(1,1,c),(i_{\max},j_{\max},c)\neq(1,1,c), \tag{11}$$

where N is the number of blocks in image,

c is the color channel,

k is the number of the block being analyzed,

W_k is the matrix of Walsh-Hadamard transformants of the original image for the k-th block,

δ the Kronecker delta, which is defined as

$$\delta\left((i,j,c),\max_{i_{\max},j_{\max},c}W_k(i_{\max},j_{\max},c)\right)=\begin{cases}1,(i,j,c)=(i_{\max},j_{\max},c),\\0,(i,j,c)\neq(i_{\max},j_{\max},c).\end{cases} \tag{12}$$

We present the experimentally obtained matrix $M(i,j,c)$ for the case of the red channel, where the values corresponding to the low-frequency Walsh-Hadamard transformants often used in practice are highlighted in blue

$$M(i,j,c)=\begin{bmatrix}0 & 0.008 & 0.03 & 0.009 & 0.25 & 0.004 & 0.07 & 0.006\\ 0.01 & 0.004 & 0.003 & 0.002 & 0.003 & 0.002 & 0.003 & 0.002\\ 0.026 & 0.003 & 0.006 & 0.003 & 0.009 & 0.002 & 0.006 & 0.0025\\ 0.1 & 0.003 & 0.004 & 0.003 & 0.005 & 0.002 & 0.003 & 0.002\\ 0.3 & 0.004 & 0.009 & 0.005 & 0.06 & 0.002 & 0.0025 & 0.003\\ 0.0055 & 0.002 & 0.002 & 0.002 & 0.003 & 0.002 & 0.002 & 0.002\\ 0.2 & 0.003 & 0.005 & 0.003 & 0.03 & 0.002 & 0.0085 & 0.0025\\ 0.008 & 0.002 & 0.002 & 0.002 & 0.004 & 0.002 & 0.003 & 0.002\end{bmatrix}. \tag{13}$$

Similarly, matrices $M(i,j,c)$ can be obtained for other color channels, in particular, for channels of the YCbCr color space, if it is used for information embedding.

Based on the obtained results (graphically presented in Fig. 1) and a series of performed experiments, three primary conditions were derived. These conditions allow for the blind detection (i.e., without access to the original image) of images affected by the steganographic method with code control.

The conditions are as follows:

1.Threshold values for indices (5,1), (5,5), and (1,5) are set at 0.3; 0.06, and 0.25, respectively. This condition is used to minimize False Negative results.

2.The analyzed value in the examined color channel must differ from the values in the other channels by at least 0.02 (experiments with higher and lower values yielded less satisfactory results). This condition is also used to minimize False Negative results.

3.At least in 40% of cases, all other values in the color channel, except for the analyzed one, must be lower than the corresponding values in the other channels (experiments with higher and lower percentages yielded less satisfactory results). This condition is used to minimize False Positive results.

Thus, the fulfillment of all three conditions is necessary to identify an image as having been affected by the steganographic method with code control. This allows for the determination of the specific channel where the information was embedded, as well as the transformant that was influenced (which indicates the type of codeword used).

Experimental results. Proposed method. A series of experiments were performed using the selected image datasets. The steganographic message sets in lossless TIFF format using the codewords affecting (5,1), (5,5), and (1,5) were analyzed in the RGB space (embedding in the red channel) and the YCbCr space (embedding in Y). The datasets in lossy JPEG format (at 100% quality) were analyzed in the RGB space (embedding in the red channel). In all cases, embedding occurred in 100% of the blocks. The results of the percentage of errors are presented in Table 2. It should be noted that False Positive cases are those where multiple embedding locations/channels are detected (even if one of them is correct) when it is known that there is only one. False Negative refers to cases where no impact was detected.

Table 2.

The number of errors when using the proposed method of steganalysis

Format	Space	Code	False Positive	False Negative	All errors
.tif	RGB	(5,1)	3,4%	16,3%	19,7%
.tif	RGB	(5,5)	7,7%	13,9%	13,9%
.tif	RGB	(1,5)	5,3%	14,7%	20%
.tif	YCbCr	(5,1)	5,3%	10,2%	15,5%
.tif	YCbCr	(5,5)	7,7%	4,6%	12,3%
.tif	YCbCr	(1,5)	5,5%	11,3%	16,8%
.jpg	RGB	(5,1)	0%	34,9%	34,9%
.jpg	RGB	(5,5)	0%	4,3%	4,3%
.jpg	RGB	(1,5)	0%	23,2%	23,2%

Table 2 shows that the best results are obtained when operating in YCbCr space in a lossless format. The worst results in two of the three cases (namely for codewords (5,1) and (1,5)) produce lossy .jpg images. However, they also give zero false positives — which is also a good result. This may be due to small changes in the average threshold results.

Another observation is that statistically, detection works best when the codeword affecting (5,5) transformant is used for embedding information, meaning that the (2,2) DCT transformant is affected. This is because the threshold value for this index is quite low (0,06), making a significant jump in the color channel more noticeable compared to other channels. As a result, this jump more frequently meets the established three conditions, contributing to better detection in most cases.

The method was also tested for different percentages of utilized blocks (for lossless .tif format, codeword affecting (5,1) transfromant, embedding in the red color channel). Table 3 contains the obtained results.

Table 3.

The number of errors when using the proposed steganalysis method depending on the percent of blocks embedded

Format	Space	% of blocks embedded	Code	False Positive	False Negative	All errors
.tif	RGB	100	(5,1)	3,4%	16,3%	19,7%
.tif	RGB	70	(5,1)	3,8%	20,4%	24,2%
.tif	RGB	50	(5,1)	3,2%	27,6%	30,8%
.tif	RGB	25	(5,1)	4,7%	47,1%	51,8%
.tif	RGB	10	(5,1)	4,5%	60,7%	65,2%

Thus, in this case, it is evident that at 25% of blocks embedded (in large images), the presented method loses effectiveness and detects the covert channel with code control in less than 50%. However, at 100-50% of blocks embedded, it still operates at a fairly significant level of correct detections.

Similar methods. In general, steganalysis methods can be categorized based on various criteria. First, they can be classified according to the information available to the analyst. For example, the steganographic object is known; the steganographic object and the container are known; the hidden message is known; the algorithm is known; both the hidden message and the algorithm are known, as well as the case where all of the aforementioned elements are known.

The proposed method is a blind method since only the steganographic object is known. Currently developing blind steganalysis methods can be categorized into statistical analysis methods [14], adaptive steganalysis methods [15, 16], and methods based on deep learning [17, 18]. These methods of steganalysis are promising, but they suffer from disadvantages such as high computational requirements and the need for large training data sets.

A popular tool that uses the methods of adaptive steganalysis is StegExpose. StegExpose is a steganalysis tool that specializes in detecting steganography in lossless images such as PNG and BMP (LSB detection methods). It has a command-line interface and is designed for batch image analysis, providing reporting capabilities and intuitive settings [19].

This tool provides blind analysis. Testing this tool on the aforementioned datasets of influenced images in lossless TIFF format led to the conclusion that it is unable to detect covert communication channels created using a steganographic method with code control of information embedding. The experiments showed that correct detection occurs only in 0.2% of cases for any codeword, while a false positive result occurs in 99.8% of cases, demonstrating the complete ineffectiveness of the utility.

The paper [20] also discusses several tools for detecting LSB embedding, including LSB-matching, but most of them require the presence of the original image, which diminishes their value for blind detection. In general, methods for detecting LSB embedding [21, 22] most often require the original image, or they analyze the images in grayscale, which makes it impossible to determine the specific embedding channel, as proposed by the method presented in this paper. Table 4 presents the results of the analysis of the aforementioned datasets in lossless TIFF format using the histogram comparison method [21].

Table 4.

The number of errors when using the method of comparative analysis of histograms

Format	Space	Code	True Positive	False Negative
.tif	RGB	(5,1)	32,3,%	67,7%
.tif	RGB	(5,5)	37,2%	62,8%
.tif	RGB	(1,5)	29,6%	70,4%

The analysis of Table 4 leads to the conclusion that the number of errors significantly exceeds the number of errors when using the proposed method.

Conclusions. In this paper, a steganalysis method for detecting a covert channel with code control is presented. This method uses the Walsh-Hadamard transform domain for analysis, which is a promising field for further research and the development of more steganalysis methods. This method is mathematically simple and offers fewer errors than other listed analogs; it also provides the ability to identify the specific transform that has been affected and the channel into which additional information was embedded using the steganographic method with code control.

Such further research directions can be highlighted: experimental determination of more effective threshold values for all other block indices, apart from those considered (5,1), (5,5), and (1,5); exploring the possibility of detecting embedding using the codeword (1,1), as this component is not accounted for the calculations. Further research is also needed to analyze images that have been attacked by compression.

The Walsh-Hadamard transform domain opens up significant potential for such research.

References

1. Chang C. C., Liu Y., Chen K. Real-time adaptive visual secret sharing with reversibility and high capacity. *J. Real-Time Image Process*. 2019. No. 16. P. 871-881. DOI: 10.1007/s11554-018-0813-9.
2. Zyara A. Suggested method for hiding secret data in cover image depending on the Connected Component Labeling algorithm. *International Journal of Electrical and Computer Engineering (IJECE).* 2016. P. 2349-7084.
3. Wang Y., Tang M., Wang Z. High-capacity adaptive steganography based on LSB and Hamming code. *Optik*. 2020. P. 164685. DOI: 10.1016/j.ijleo.2020.164685.
4. Ker A. Improved detection of LSB steganography in grayscale images. Information Hiding Workshop. 2004. V. 3200. P. 97-115.
5. Zhiqiang Z., Ning Z., Tong Q., Ming X. Robust Steganography by Modifying Sign of DCT Coefficients. *IEEE Access.* 2019. P. 1-16. DOI: 10.1109/ACCESS.2019.2953504.
6. Di F., Zhang M., Huang F., Liu J., Kong Y. Reversible data hiding in JPEG images based on zero coefficients and distortion cost function. *Multimedia Tools Applications.* 2019. V. 78. P. 34541–34561. DOI: 10.1007/s11042-019-08109-8.
7. Ping P., Zeming W., Bing Y. C., Bing Z. Double-Matrix Decomposition Image Steganography Scheme Based on Wavelet Transform with Multi-Region Coverage. *Entropy.* 2022. V. 24. P. 256. DOI: 10.3390/e24020246.
8. Valandar M. Y., Ayubi P., Barani M. J. A new transform domain steganography based on modified logistic chaotic map for color images. Journal of Information Security and Applications. 2017. No. 3. P. 142-151. doi: 10.1016/j.jisa.2017.04.004.
9. Hamidi M., Haziti M. E., Cherifi H., Hassouni M. E. Hybrid blind robust image watermarking technique based on DFT-DCT and Arnold transform. Multimedia Tools Applications. 2018. Vol. 77. P. 1-34. doi: 10.48550/arXiv.1911.00753.
10. Kobozeva A., Sokolov A. Robust Steganographic Method with Code-Controlled Information Embedding. Problems of the Regional Energetics. 2021. P. 115-130. Vol. 4. doi: 10.52254/1857-0070.2021.4-52.11.
11. Kobozeva A., Sokolov A. Theoretical foundations for constructing effective codewords for the code-controlled information embedding steganographic method. *Radiotekhnika.* 2021. No. 4. P. 27-39. DOI: 10.30837/rt.2021.4.207.02.
12. Bhattacharyya S. A Robust Image Steganography using Hadamard Transform. *International Conference on Information Technology in Signal and Image Processing*. 2013. P. 132-142.
13. Steganalysis: How to Detect Steganography, [Electronic resource]. 2018. URL: https://digitnet.github.io/m4jpeg/about-steganography/how-to-detect-steganography.htm.

14. Cai K., Li X., Zeng T., Yang B., Lu X. Reliable histogram features for detecting LSB matching. *IEEE International Conference on Image Processing, Hong Kong, China*. 2010. P. 1761 – 1764. DOI: 10.1109/ICIP.2010.5651567.
15. Jackson J. T., Gunsch G. H., Claypoole R. L., Lamont G. B. Blind Steganography Detection Using a Computational Immune System: A Work in Progress. *International Journal of Digital Evidence*. 2003. V. 4. P. 1-19.
16. StegAlyzerAS. 2018. URL: https://www.sciencedirect.com/topics/computer-science/steganography-tool.
17. Lin J., Yang Y. Multi-Frequency Residual Convolutional Neural Network for Steganalysis of Color Images. *IEEE Access*. 2021. No. 9. P. 1-13. DOI: 10.1109/ACCESS.2021.3119664.
18. Agarwal S., Kim C., Jung K.-H. Steganalysis of Context-Aware Image Steganography Techniques Using Convolutional Neural Network. *Appl. Sci.* 2022. No. 12. P. 10793. DOI: 10.3390/app122110793.
19. StegExpose. 2015. URL: https://github.com/b3dk7/StegExpose.
20. Pelosi M., Easttom C. Identification of LSB image Steganography using Cover Image Comparisons. *Journal of Digital Forensics Security and Law*. 2021. No. 15. P. 6. DOI: 10.15394/jdfsl.2021.1551.
21. Jung K. H. Comparative Histogram Analysis of LSB-based Image Steganography. *WSEAS Transactions on Systems and Control*. 2018. No. 13. P.1991-8763.

СТЕГАНОАНАЛІЗ МЕТОДУ ІЗ КОДОВИМ УПРАВЛІННЯМ ВБУДОВУВАННЯ ІНФОРМАЦІЇ В ОБЛАСТІ ПЕРЕВТОРЕННЯ УОЛША-АДАМАРА

О. О. Лановська, А. В. Соколов

Національний університет «Одеська політехніка»
1, Шевченка пр., м.Одеса, 65044, Україна
Email: radiosquid@gmail.com

Стеганографія є невід'ємною частиною побудови сучасних систем захисту інформації. Стеганографічний метод з кодовим управлінням вбудовуванням інформації є сучасним стеганографічним методом, що оперує в просторовій області контейнеру. Перевагами цього методу є забезпечення надійності сприйняття, стійкості до атак проти вбудованого повідомлення, достатньої пропускної спроможності, високої обчислювальної ефективності. На противагу стеганографічним методам створюються методи стеганоаналізу, що дозволяють виявляти вбудовану інформацію. На сьогоднішній день дослідження щодо стійкості стеганографічного методу з кодовим управлінням до атак криптоаналізу не проводилися, специфічні методи виявлення втручання за допомогою стеганографічного методу з кодовим управлінням невідомі. Метою цієї роботи є розробка методу стеганоаналізу для виявлення прихованого каналу, організованого за допомогою стеганографічного методу з кодовим управлінням вбудовування інформації в області перетворення Уолша-Адамара. У роботі проведені дослідження поведінки трансформант перетворення Уолша-Адамара контейнерів та стеганоповідомлень, які дозволили сформулювати умови наявності додаткової інформації в зображенні. Зазначені умови стали основою для розробки легкого та математично простого методу стеганоаналізу, який використовує область перетворення Уолша-Адамара. Проведені дослідження запропонованого методу дозволили встановити його високу ефективність в різних умовах при низький обчислювальній складності. Зокрема показано, що ефективність запропонованого методу перевищує ефективність методу аналізу гістограм, та методів, реалізованих у відомому інструменті StegExpose. Отримані результати дозволяють рекомендувати запропонований метод стеганоаналізу для практичного застосування для виявлення прихованих каналів зв'язку, що створенні із застосуванням стеганографічного методу з кодовим управлінням. Зокрема, зважаючи на простоту своєї алгоритмічної реалізації, запропонований метод буде ефективним в умовах обмежених обчислювальних ресурсів.
Ключові слова: стеганографія, перетворення Уолша-Адамара, стеганоаналіз, кодове управління.

DOI 10.15276/imms.v14.no4.284
UDK 681.5:621.18-5

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 284-295

MODEL OF GREENHOUSE GAS EMISSION MINIMIZATION UNDER VARIABLE LOAD OF A STEAM BOILER

T. D. Markolenko, D. G. Prodanov

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: tanyadmb@ukr.net

The article investigates the problem of reducing greenhouse gas emissions during variable load of steam generating energy equipment. The aim of the research is to develop a method and model that allow reducing greenhouse gas emissions by using a mixture of regulated methane and alternative low-calorific gas. The scientific and practical significance of the work lies in the creation of regulation methods that ensure the ecological and energy efficiency of the equipment without its modernization. The research methodology is based on mathematical modeling of the combustion process of a gas mixture using chemical kinetics equations for the isoenthalpic process. A model has been developed within the study that allows accurately determining the adiabatic combustion temperature of a fuel-air mixture, as well as the quantitative and qualitative composition of the combustion products formed under different ratios of methane and low-calorific gas. A relationship between the fuel mixture consumption, its composition, and the stability of the combustion products volume has been identified. It has been shown that maintaining constant exhaust gas flow can be achieved by adjusting the fuel mixture composition, which influences the equipment power and the temperature of the exhaust gases. The effect of maintaining isoenthalpic process conditions on the amount of energy released has been separately considered, which allowed determining the patterns for regulating the equipment power without significant losses in energy efficiency. The value of the study lies in the development of new approaches to energy efficiency management and greenhouse gas emissions reduction. The practical significance lies in the possibility of applying the obtained results to regulate equipment power without modernization, while maintaining a constant volume of combustion products.
Keywords: power regulation, low-calorific gas, greenhouse gases, chemical kinetics, isoenthalpic process.

Introduction. One of the directions of the international decarbonization program for industry and energy is the development of renewable energy sources. "Green" energy requires significant capital investments. Therefore, it is fully accessible only to a limited number of industrially developed countries. However, even these countries, in order to ensure the dispatchability of the energy supply, retain the majority of their generation based on the combustion of extractable fuels. Furthermore, the preservation and even development of traditional energy generation sources is even more pronounced in countries with limited financial resources.

Dispatchability is determined by the ability to make operational changes in the power of the energy equipment being operated within a wide range. At the same time, it remains crucial to address the issue of minimizing greenhouse gas emissions across the entire operational range of traditional hydrocarbon energy equipment based on extracted fuels.

Analysis of research and publications. The elimination or significant reduction of greenhouse gas emissions in the future is considered from the perspective of transitioning to hydrogen as a fuel. Even at the stage of analyzing prospects, a number of issues arise during the development of new technology [1]. Significant financial investments are required to create hydrogen production technologies, develop the necessary logistics infrastructure, and implement energy equipment.

Currently, or in the near future, instead of fully hydrogen-based energy, the following possible measures can be identified for reducing greenhouse gas emissions:

– mixing traditional hydrocarbon fuels with hydrogen in a specific proportion, which does not require a radical change in the equipment used;

– combustion of organic material of non-extracted origin with a specified energy production volume to reduce greenhouse gas emissions, accompanied by a special technological regulation of the energy installation.

The current state of hydrogen production technology is primarily determined by the reforming of extracted organic fuel [2]. By 2021, out of 60 million tons of hydrogen produced annually, approximately 96% was obtained by processing extracted fuels (49% natural gas, 29% liquid hydrocarbons, and 18% coal), with only 4% produced by water electrolysis. This situation resulted in high indirect carbon dioxide emissions [3]. It should be noted that the energy for hydrogen production processes largely depends on the extracted organic fuel used. Thus, using hydrogen as a fuel in industry and energy will lead to higher carbon dioxide emissions than using fossil fuels for the same energy output.

In the absence of developed technology for producing carbon-neutral "green hydrogen from renewable energy sources" and "orange hydrogen from nuclear energy," the use of its "gray" variant may currently be justified in research and demonstration projects assessing the environmental impact. It should be noted that a series of studies has been conducted that allows for increasing the reliability of fuel rod operation in various modes [4, 5], which makes it possible to consider nuclear power plants as hydrogen suppliers. There are known studies on automated control models and methods for adjusting power in nuclear power plants to ensure stable control in load-following modes, which effectively allows for regulating hydrogen oxidation [6, 7].

In [8, 9], the issue of safety regarding the use of a methane-hydrogen mixture in existing boiler equipment is addressed. In [8], the moderate impact of hydrogen content up to 50% in the fuel mixture on the energy and environmental characteristics of the equipment is noted, and the possibility of safe operation at such concentrations is demonstrated. In [9], a hydrogen concentration of 50% in the fuel mixture in the pipeline is considered moderately hazardous, while a concentration up to 25% is regarded as safe. It should be noted that this refers to the safety of the mixture during normal equipment operation (without leaks). While the possibility of a gas fuel mixture leak can be minimized, it cannot be entirely ruled out. The methane-hydrogen mixture with air is more explosive than pure methane due to hydrogen's lower minimum ignition energy in air compared to methane (0,020 mJ versus 0,29 mJ – an order of magnitude difference) [1, Table 1].

There is also a thermophysical feature of the methane-hydrogen mixture that complicates its use in existing equipment. The gravimetric calorific value of hydrogen is higher compared to methane (119.9 MJ/kg versus 45.8 MJ/kg), which gives hydrogen an advantage when transported in liquid form. However, hydrogen is supplied to energy equipment in a gaseous state. In this case, the volumetric energy ratio is reversed (10.7 MJ/m^3 versus 33 MJ/m^3) [1, Table 1]. It is important to consider the potential decrease in the power output of installed energy equipment as the hydrogen concentration in the fuel mixture increases, in the absence of adjustments to the geometric dimensions of the gas supply and exhaust fuel pathways.

Boilers that use natural gas as fuel have an efficiency of 88-93%. Improving this efficiency reduces carbon dioxide emissions for a given energy production volume. In the energy balance of boilers, the largest losses occur through flue gases (5-10%). Therefore, the main area where significant efficiency improvements can be achieved is by reducing these losses. For example, in [10], a method was developed to determine the variable composition and energy content (enthalpy) of gaseous fuel during combustion. This method enables the calculation of boiler efficiency based on the current load and flue gas temperature.

A change in the load of a gas boiler causes a corresponding change in the flue gas temperature. The need for cyclic power adjustments in a non-condensing boiler within the range of 40% to 100% [11] can result in flue gas temperatures varying between ~393 K and 473 K (120°C to 200°C). Maximum load corresponds to the highest temperature and the greatest

losses. The minimum temperature is determined by the need to maintain a non-condensing mode during flue gas evacuation. The presence of hydrogen in the fuel mixture, in quantities deemed safe for operational conditions, slightly increases [12, Table 4] the dew point temperature of the flue gases. This phenomenon is advantageous for condensing boilers [13] but undesirable for non-condensing energy systems [14]. In such systems, to prevent condensate formation in the flue gas channels, the flue gas temperature must be increased, which leads to a corresponding decrease in efficiency.

Theoretically, in a non-condensing flue gas removal mode, an efficiency improvement of up to ~4% could be achieved by lowering the maximum flue gas temperature to an acceptably low level. However, this approach is not feasible with the current designs of energy equipment. It should also be noted that, in a cyclic variable mode, the energy system operates at maximum load for only a small fraction of the time. Consequently, even in theory, the average efficiency improvement would amount to less than 4%.

Greenhouse gases are formed during the use of fossil fuels. They do not include alternative combustible gases obtained from the processing of secondary raw materials. Examples of such gases are the products of wood gasification or agricultural waste [15]. These gases are characterized by a lower calorific value compared to methane. Therefore, a complete replacement of methane with alternative combustible gases would lead to a decrease in the maximum capacity of the installed equipment.

The use of a fuel mixture of methane and alternative combustible gases obtained from organic raw materials or renewable sources, differentiated according to the specified load, appears to be relevant in energy steam-generating installations.

Research Objective. The objective of the research is to develop a method and model for reducing greenhouse gas emissions under varying load conditions of steam-generating energy equipment through the use of a regulated mixture of methane and alternative gas.

To achieve this goal, the following tasks were set:

– to develop a method and model for forming the composition of the combustible gas mixture that corresponds to the specified load of the steam-generating energy equipment;

– to determine the impact of the mixture composition on the parameters of the combustion products (flue gases);

– to determine the influence of the technical parameters of the installed energy equipment on the composition of the combustible gas mixture while ensuring the specified operating mode.

Main Part. The use of a low-calorific component in the combustible gas mixture can have different effects on the operation of the installed equipment. On one hand, it will lead to a change in the quantity of flue gases and, accordingly, affect the equipment's efficiency. On the other hand, reducing the calorific value of the gas fuel mixture increases the required volume to ensure the specified power output. This feature must be considered when evaluating the throughput capacity of the fuel supply system.

With varying boiler load, the adiabatic combustion temperature of the fuel remains unchanged if a constant gas composition is used and optimal excess air is provided. However, the temperature of the flue gases is variable and reaches its maximum at maximum load. The reason for this is well known. At maximum load, the greatest amount of combustion products is formed. These pass through the gas ducts at maximum speed, and, accordingly, spend the least amount of time in contact with the heat exchange surfaces. To account for this feature, one of the methods for regulating the flue gas temperature and, as a result, the boiler's efficiency, may be applied.

It can be assumed that the addition of ballast gases will influence the change in the temperature of the flue gases. For instance, as shown in [16, Table 5], at a lower load (81% versus 99%), but with a correspondingly higher air excess ratio (1.57 versus 1.2), the temperature of the flue gases was higher (167°C versus 115°C). This occurs when using fuel of constant composition. When the optimal amount of air for combustion is supplied, the

temperature ratio should be reversed. This effect can be explained by the air excess acting as a ballast gas. A similar effect can be observed when fuel is supplied as a mixture of gases, as additionally shown in [17]. Preliminary studies [18] and further development of models and methods [19] have made it possible to consider various options for using natural (methane) and blast furnace gases in different proportions by developing the fundamentals for conducting such processes [20]. In [21], the option of using natural (methane) and blast furnace gases in different proportions is discussed. The latter has a deliberately lower calorific value and contains several ballast gases. It is worth noting that at the same load (for example, 40 tons per hour of steam), when the share of blast furnace gas in the fuel mixture increases from 0.198 to 0.755, the temperature in the furnace decreases, and the temperature of the flue gases increases from 148°C to 174°C [21, Table 2].

Thus, there is a potential opportunity to control the temperature of the exhaust gases at a given boiler load. To increase the temperature, it is necessary to replace some portion of the standard fuel (e.g., methane) with a gas of lower calorific value, in a corresponding greater amount. Controlling the exhaust gas temperature allows it to remain constant across the entire load range. However, this results in an increase in temperature to the value corresponding to the maximum load. Afterward, it is necessary to design a low-temperature economizer to reduce the exhaust gas temperature to levels that prevent the condensation of water vapor in the flue gas system.

The implementation of such a steam generation method and control strategy may require the use of a large volume of low-calorific fuel gas mixtures. The possible volume could be limited by the throughput capacity of the existing fuel mixture supply system and may require its modernization.

Using only the existing capacity of the fuel gas mixture delivery line presents another potential opportunity to reduce greenhouse gas emissions at loads different from the maximum. When using standard gas, such as methane, reducing the load is accompanied by a decrease in its supply. The amount of methane supplied can be further reduced by mixing in a low-calorific gas to the volume corresponding to the capacity of the fuel delivery line, while maintaining the desired load. When implementing this management method, it is essential to monitor the amount and composition of the flue gases produced. If their volume is too low, there is a risk of entering an undesirable condensation operation mode.

The presented analysis of the two control methods allows for the proposal of an operational method for a steam-generating energy unit that enables controlling (or setting) the composition of the fuel gas mixture.

The method for calculating the composition of the gas mixture. In the case of combustion process equilibrium, the composition and quantity of flue gases can be determined using a balance method. However, a more universal approach is to use the calculation method based on chemical kinetics equations. In this case, it is possible to determine not only the temperature of the combustion products but also their composition, if it depends on this temperature. For example, this applies to nitrogen oxides. Moreover, the calculation model can be structured in such a way that the composition of the flue gases is determined for the same volume but with different fuel mixture compositions. At the same time, the required volume of the fuel mixture will also be determined.

The calculation method is as follows:

– the volumetric fraction of the main gas is gradually reduced (from 100% to 0%) and is replaced with the corresponding amount of lower-calorific gas. For each of these ratios, knowing the compositions of the input gases, air, and the excess air coefficient α, the composition of the fuel-air mixture and its enthalpy are determined;

– using a model based on chemical kinetics equations, the composition of flue gases (in volumetric fractions) is determined for each type of fuel-air mixture, depending on the amount of lower-calorific gas;

– each individual composition of flue gases is calculated for the same volume when using any type of fuel-air mixture, for example, for 1 mole of the burning substance. A key feature of the method is the calculation of the required amount (M_T) of the corresponding fuel-air mixture in moles to produce 1 mole of flue gases. Since both the initial fuel-air mixture and combustion products are in the gaseous state, all mole ratios correspond to volumetric ones;

– for each fuel-air mixture composition, the amount of energy Q_m transferred by 1 mole of flue gases when cooled to a specified temperature (e.g., 393 K or 120°C) is determined. This is necessary for further application regarding the ratio of standard gas to lower-calorific-value gas at different boiler loads. Additionally, the lower heating value relative to 1 mole of flue gases can be determined, and if desired, the lower heating value can be converted using the M_T for 1 mole of the fuel-air mixture. The obtained results can be used to determine the ratio of standard gas to lower-calorific-value gas, as well as the required amount of the resulting mixture to ensure the boiler operates at the specified load. This ensures the consistency of the amount of flue gases.

Model. The input data are determined by the gross formula of the fuel mixture. The combustible gases used in energy production have a hydrocarbon composition and consist of a limited number of chemical elements. Taking into account the composition of the air used as an oxidizer, the list may include C, H, O, and N. In some cases, sulfur (S) may be present, but its presence or absence does not affect the generality of the solution. The gross formula of the fuel mixture (excluding sulfur) will have the following form:

$$\mathrm{C}_{b_\mathrm{C}}\mathrm{H}_{b_\mathrm{H}}\mathrm{O}_{b_\mathrm{O}}\mathrm{N}_{b_\mathrm{N}}, \tag{1}$$

where b_C, b_H, b_O and b_N are the number of atoms of the respective elements in the gross formula. Their values are determined based on the ratio of the primary and lower-calorific gases, as well as the excess oxidizer coefficient. Additionally, based on the enthalpy values of the input gases and air, the specific enthalpy of the fuel-air mixture is determined (1).

Based on equation (1), the list of substances that may be present in the flue gases is determined. If the oxidizer excess coefficient $\alpha > 1$, and the combustion process is in equilibrium, this list may include:

$$\mathrm{CO_2},\ \mathrm{H_2O},\ \mathrm{O_2},\ \mathrm{N_2},\ \mathrm{NO}. \tag{2}$$

In the case of $\alpha < 1$, gases that are formed due to incomplete combustion of the fuel mixture may appear: CO, $\mathrm{H_2}$, $\mathrm{CH_4}$. In this case, we will limit ourselves to the list in equation (2).

In the equilibrium process of thermodegradation of the fuel mixture, for substances from the list (2), the chemical kinetics equation can be written according to the law of mass action. Given that the substance exists in the gas phase, it is more convenient to express this equation using partial pressures rather than molar concentrations. In this case, the calculations should take into account not just one mole of the initial raw material (along with the oxidizer), but a certain number of its moles (M_T), which leads to a numerical equivalence between molar concentration and partial pressure. This value is determined during the calculations. For the products in (2), the law of mass action equation will have the following form:

$$\frac{P_\mathrm{C}\cdot P_\mathrm{O}^2}{P_{\mathrm{CO_2}}} = K_{\mathrm{CO_2}}(T);\ \frac{P_\mathrm{H}^2\cdot P_\mathrm{O}}{P_{\mathrm{H_2O}}} = K_{\mathrm{H_2O}}(T); \tag{3}$$

$$\frac{P_\mathrm{O}^2}{P_{\mathrm{O_2}}} = K_{\mathrm{O_2}}(T);\ \frac{P_\mathrm{N}^2}{P_{\mathrm{N_2}}} = K_{\mathrm{N_2}}(T);\ \frac{P_\mathrm{N}\cdot P_\mathrm{O}}{P_{\mathrm{NO}}} = K_{\mathrm{NO}}(T). \tag{4}$$

In the denominators, there are the partial pressures of the corresponding substances. In the numerators, there are the partial pressures of the atoms of the chemical elements that make up these substances. The exponent is equal to the number of corresponding atoms in the substance's formula. $K_{\mathrm{CO_2}}(T)$; $K_{\mathrm{H_2O}}(T)$; $K_{\mathrm{O_2}}(T)$; $K_{\mathrm{N_2}}(T)$; $K_{\mathrm{NO}}(T)$ – the equilibrium constants of the corresponding substances at the current combustion temperature (tabulated values).

Using (1), material balance equations for each chemical element included in this formula can be written for the desired model:

for [C] $$b_{\mathrm{C}} \cdot M_T = P_{\mathrm{CO_2}} + P_{\mathrm{C}}; \quad (5)$$

for [H] $$b_{\mathrm{H}} \cdot M_T = 2 \cdot P_{\mathrm{H_2O}} + P_{\mathrm{H}}; \quad (6)$$

for [O] $$b_{\mathrm{O}} \cdot M_T = 2 \cdot P_{\mathrm{CO_2}} + P_{\mathrm{H_2O}} + 2 \cdot P_{\mathrm{O_2}} + P_{\mathrm{NO}} + P_{\mathrm{O}}; \quad (7)$$

for [N] $$b_{\mathrm{N}} \cdot M_T = 2 \cdot P_{\mathrm{N_2}} + P_{\mathrm{NO}} + P_{\mathrm{N}}. \quad (8)$$

In writing these equations, it is taken into account that not one, but M_T moles of the initial fuel-air mixture with the gross formula (1) are considered.

To close the system, Dalton's law is used:

$$P_{\Sigma} = P_{\mathrm{CO_2}} + P_{\mathrm{H_2O}} + P_{\mathrm{O_2}} + P_{\mathrm{N_2}} + P_{\mathrm{NO}} + P_{\mathrm{C}} + P_{\mathrm{H}} + P_{\mathrm{O}} + P_{\mathrm{N}}. \quad (9)$$

Here, P_{Σ} is the pressure inside the boiler furnace.

Equations (3–9) form a closed nonlinear algebraic system for determining 10 values – 9 partial pressures and M_T.

$$P_{\mathrm{CO_2}}, P_{\mathrm{H_2O}}, P_{\mathrm{O_2}}, P_{\mathrm{N_2}}, P_{\mathrm{NO}}, P_{\mathrm{C}}, P_{\mathrm{H}}, P_{\mathrm{O}}, P_{\mathrm{N}}, M_T. \quad (10)$$

Features of the solution. The solution of the obtained system is complicated by the large ratio of the values of the variables included in it. This ratio can exceed 15 orders of magnitude ($>10^{15}$). To overcome this problem, all the equations in the system can be logarithmically transformed. During the solution process, the values of pressures and M_T, which are logarithms, are determined. In this case, the resulting values will differ by factors, not orders of magnitude.

The original system (3-9) will take the form:

$$\ln\left(P_{\mathrm{C}}\right) + 2 \cdot \ln\left(P_{\mathrm{O}}\right) - \ln\left(P_{\mathrm{CO_2}}\right) = \ln\left(K_{\mathrm{CO_2}}(T)\right); \quad (11)$$

$$2 \cdot \ln\left(P_{\mathrm{H}}\right) + \ln\left(P_{\mathrm{O}}\right) - \ln\left(P_{\mathrm{H_2O}}\right) = \ln\left(K_{\mathrm{H_2O}}(T)\right); \quad (12)$$

$$2 \cdot \ln\left(P_{\mathrm{O}}\right) - \ln\left(P_{\mathrm{O_2}}\right) = \ln\left(K_{\mathrm{O_2}}(T)\right); \quad (13)$$

$$2 \cdot \ln\left(P_{\mathrm{N}}\right) - \ln\left(P_{\mathrm{N_2}}\right) = \ln\left(K_{\mathrm{N_2}}(T)\right); \quad (14)$$

$$\ln\left(P_{\mathrm{N}}\right) - \ln\left(P_{\mathrm{O}}\right) - \ln\left(P_{\mathrm{NO}}\right) = \ln\left(K_{\mathrm{NO}}(T)\right); \quad (15)$$

$$\ln\left(b_{\mathrm{C}}\right) + \ln\left(M_T\right) = \ln\left(P_{\mathrm{CO_2}} + P_{\mathrm{C}}\right); \quad (16)$$

$$\ln\left(b_{\mathrm{H}}\right) + \ln\left(M_T\right) = \ln\left(2 \cdot P_{\mathrm{H_2O}} + P_{\mathrm{H}}\right); \quad (17)$$

$$\ln\left(b_{\mathrm{O}}\right) + \ln\left(M_T\right) = \ln\left(2 \cdot P_{\mathrm{CO_2}} + P_{\mathrm{H_2O}} + 2 \cdot P_{\mathrm{O_2}} + P_{\mathrm{NO}} + P_{\mathrm{O}}\right); \quad (18)$$

$$\ln\left(b_{\mathrm{N}}\right) + \ln\left(M_T\right) = \ln\left(2 \cdot P_{\mathrm{N_2}} + P_{\mathrm{NO}} + P_{\mathrm{N}}\right); \quad (19)$$

$$\ln(P_{\Sigma}) = \ln(P_{\mathrm{CO_2}} + P_{\mathrm{H_2O}} + P_{\mathrm{O_2}} + P_{\mathrm{N_2}} + P_{\mathrm{NO}} + P_{\mathrm{C}} + P_{\mathrm{H}} + P_{\mathrm{O}} + P_{\mathrm{N}}). \quad (20)$$

The combustion process in the boiler is considered isoenthalpic. The calculation method is based on determining, through an iterative process, the temperature and the corresponding composition of the reaction products (flue gases) so that their total enthalpy equals the enthalpy of the incoming fuel-air mixture.

Results. The solution is obtained based on the method and model presented above (3-9).

Let's represent the gross formula of the fuel-air mixture. Methane $\mathrm{CH_4}$ is taken as the standard fuel gas, with a formation enthalpy of $I_{\mathrm{CH_4}} = -74.85$ kJ/mol (298 K). The low-calorific gas is a mixture of gases produced during the decomposition of pine wood waste in a gasifier. In the considered case, the cooled gas mixture is at 298 K, and after the removal of condensate,

its molar composition is as follows: $CO-0.394$, $CO_2-0.214$, $H_2-0.360$, $H_2O-0.032$ with the gross formula:

$$C_{0.61}H_{0.78}O_{0.85} \tag{21}$$

and the enthalpy of formation $I_{mk}=-135.5$ kJ/mol. The absence of nitrogen is due to the fact that the gasifier, during the decomposition of wood waste, is purged not with air but with oxygen [10]. Such processes are now implemented on an industrial scale.

We define the share of low-calorific gas (21) in the mixture with methane as $\varphi\in[0,1]$. When $\varphi=0$, the fuel gas consists of methane, and when $\varphi=1$, it consists entirely of gas (21). In this case, the gross formula of the fuel gas mixture at different values of will have the following form:

$$C_{1-0.39\cdot\varphi}H_{4-3.22\cdot\varphi}O_{0.85\cdot\varphi}. \tag{22}$$

The enthalpy of such a mixture is determined using the following equation:

$$I_{mix}=(1-\varphi)\cdot I_{CH_4}+\varphi\cdot I_{mk}. \tag{23}$$

Let the composition of air in molar fractions be: $O_2-0.21, N_2-0.79$. In this case, the gross formula of air will be:

$$O_{0.42}N_{1.58}. \tag{24}$$

The complete oxidation reaction during the combustion of the fuel gas mixture (22), taking into account (24), can be written as:

$$C_{1-0.39\cdot\varphi}H_{4-3.22\cdot\varphi}O_{0.85\cdot\varphi}+\left(\frac{4-3.24\cdot\varphi}{0.42}\right)O_{0.42}N_{1.58}=(1-0.39\cdot\varphi)\ CO_2+$$
$$+(2-1.61\cdot\varphi)\ H_2O+(7.52-6.05\cdot\varphi)\ N_2. \tag{25}$$

Formula (25) is written for 1 mole of the combustible gas mixture (22). In this case, the coefficient in front of the air's gross formula on the left side of the equation is the mole stoichiometric coefficient:

$$\chi_0=\frac{4-3.24\cdot\varphi}{0.42}. \tag{26}$$

Taking into account the oxidizer excess coefficient α, it shows how many moles of air are used to burn 1 mole of the combustible gas mixture (22):

$$n=\alpha\cdot\chi_0=\alpha\cdot\left(\frac{4-3.24\cdot\varphi}{0.42}\right). \tag{27}$$

Formulas (22), (24), and expression (27) allow us to determine the gross formula (1 mole) of the fuel-air mixture (1):

$$C_{\left(\frac{1-0.39\cdot\varphi}{1+n}\right)}H_{\left(\frac{4-3.22\cdot\varphi}{1+n}\right)}O_{\left(\frac{0.85\cdot\varphi+\alpha\cdot(4-3.24\cdot\varphi)}{1+n}\right)}N_{\left(\frac{\alpha\cdot(15.05-12.9\cdot\varphi)}{1+n}\right)}. \tag{28}$$

Considering air as a mixture of simple gases (O_2+N_2), we assume that its enthalpy of formation is zero. In this case, the enthalpy of formation of 1 mole of the fuel-air mixture, taking into account (23), can be determined using the relationship:

$$I=\frac{(1-\varphi)\cdot I_{CH_4}+\varphi\cdot I_{mk}}{1+n}. \tag{29}$$

The calculation results for the first regulation method with $\alpha=1.1$ are presented in Table 1.

The following abbreviations are presented in Table 1: M_{MG} – the number of moles of the combustible gas mixture (taking φ into account) required to produce 1 mole of flue gases; M_{Air} – the number of moles of air needed to burn the corresponding amount of M_{MG} moles of the combustible gas mixture; T_{ad} – the adiabatic combustion temperature; Q_{CP}^{l} – the amount of energy in the form of heat released when 1 mole of flue gases (combustion products) is cooled

to 473 K; m – is the volume multiplier of the corresponding gas mixture relative to the volume of methane required to produce 1 mole of flue gases.

Table 1.

The amount of combustible gas mixture for the same volume of flue gases

φ	M_{MG}	M_{Air}	m	Q^{l}_{CP} [kJ] 200 ºC (473 K)	T_{ad} , [K]
0	0.087	0.913	1	55.4	1965
0.1	0.095	0.912	1.09	55.8	1972
0.2	0.104	0.912	1.19	56.3	1980
0.3	0.115	0.911	1.32	56.9	1990
0.4	0.129	0.911	1.48	57.5	2000
0.5	0.146	0.910	1.68	58.5	2015
0.6	0.169	0.909	1.94	59.6	2034
0.7	0.200	0.908	2.30	61.3	2060
0.8	0.246	0.906	2.82	63.6	2095
0.9	0.318	0.902	3.65	67.4	2150
1	0.450	0.896	5.18	74.2	2243

Let us analyze the obtained results. It could have been assumed that an increase in the proportion of low-calorific gas (with an increase in φ) would lead to a decrease in the flue gas temperature and the energy released during their cooling. However, the calculation results demonstrate the opposite. As the proportion of low-calorific gas in the combustible gas mixture M_{MG} increases, the amount of air M_{Air} required for combustion remains virtually unchanged. This corresponds to a reduction in the proportion of air in the fuel-air mixture and, consequently, a decrease in the proportion of nitrogen (an inert component) in the flue gases. As a result, an increase in the proportion of low-calorific gas leads to a rise in the adiabatic combustion temperature and the amount of energy released per mole of flue gases. It is also worth noting a significant increase in the volume of the combustible gas mixture M_{MG} as the proportion of the low-calorific component grows. Consequently, with the same amount of flue gases, the complete replacement of methane with low-calorific gas results in a volume increase of m = 5.18 times. In other words, on the one hand, with the current equipment configuration, it is impossible to implement the proposed control method. The existing gas supply system cannot provide the required fuel flow rate. On the other hand, modernizing the gas supply system to increase its capacity would enable a 30% increase in the output power of the existing primary equipment. Ensuring the same volumetric flow rate of the combustible gas mixture at different coefficients constitutes the second variant of the method for controlling the equipment's operation. The calculation results are presented in Table 2. Here, as in the first variant $\alpha = 1.1$.

Table 2.

The amount of flue gases at the same volumetric flow rate of combustible gases

φ	Q_{fg} [kJ] 120 ºC (393 K)	ε	V [mole]
0	57.9	1	1
0.1	53.6	0.93	0.92
0.2	49.3	0.85	0.84
0.3	45.1	0.78	0.76
0.4	40.7	0.7	0.68
0.5	36.4	0.63	0.6
0.6	32.1	0.55	0.52
0.7	27.8	0.48	0.44
0.8	23.5	0.41	0.35

φ	Q_{fg} [kJ] 120 ºC (393 K)	ε	V [mole]
0.9	19.2	0.33	0.27
1	14.8	0.26	0.19

Here, Q_{fg} represents the amount of energy in the form of heat released when the flue gases are cooled to a temperature of 393 K. These gases are formed during the combustion of a combustible gas mixture of the same volume (this volume is equal to the volume of methane consumed to produce 1 mole of flue gases); ε denotes the fraction of the released heat relative to the amount generated during methane combustion; V represents the volume of flue gases produced during the combustion of a combustible gas mixture of the same volume (this volume is equal to the volume of methane consumed to produce 1 mole of flue gases).

In the case of using a gas of constant composition, the power reduction of the equipment is achieved by reducing the gas flow. In the first approximation, the power reduction and the decrease in gas flow are proportional. The change in the amount of flue gases happens in the same proportion. From Table 2, it follows that even when using a mixture of methane and low-calorific gas, the change in power Q_{fg} (or ε) and the volume of flue gases occurs in a similar proportion. However, the amount of methane used does not correspond to this change. For example, reducing the power to ~50% (ε=0.48) occurs when using a mixture with φ=0.7. In this case, the mixture consists of only 30% methane, instead of 50% when pure methane is used. Thus, the volume of carbon dioxide formed from the methane component will be 20% less than when pure methane is used under the same conditions. This part of the carbon dioxide is considered greenhouse gas, as it is formed from fossil fuel.

Table 3 presents the composition of the flue gases in molar fractions for different compositions of the combustible gas mixture.

Table 3.

Composition of flue gases (mole fractions)

φ	P_{CO_2}	P_{H_2O}	P_{O_2}	P_{N_2}	P_{NO}
0	0.087	0.174	0.017	0.720	0.0018
0.1	0.091	0.174	0.016	0.716	0.0018
0.2	0.096	0.174	0.016	0.712	0.0019
0.3	0.101	0.174	0.016	0.706	0.0019
0.4	0.109	0.174	0.016	0.699	0.0019
0.5	0.118	0.174	0.016	0.690	0.0020
0.6	0.129	0.175	0.016	0.678	0.0021
0.7	0.145	0.175	0.016	0.661	0.0022
0.8	0.169	0.175	0.016	0.638	0.0024
0.9	0.206	0.175	0.016	0.600	0.0026
1	0.275	0.175	0.015	0.531	0.0030

The data in Table 3 show that the mole fraction and, accordingly, the partial pressure of water vapor are almost the same for all values of φ. Therefore, the composition of the combustible gas mixture has little effect on the dew point temperature of the flue gases. The adequacy of the obtained results is confirmed by data from [20] in the calculations of pyrolysis gas.

The characteristics of the flue gases presented in Tables 2 and 3 allow us to conclude that it is possible to control the power of the installed equipment without modernization using the second method. In this case, in the maneuvering mode, as the power decreases and the proportion of low-calorific gas increases, the amount of carbon dioxide related to the greenhouse effect decreases.

The increase in the adiabatic combustion temperature when adding low-calorific gas to the mixture leads to a slight increase in the formation of nitrogen oxides compared to using only methane (Table 3). As the calculation results show, in the power regulation range (with the increase in the share of low-calorific gas to φ~0.7), the increase in the mole fraction of nitrogen oxides is negligible.

Conclusions. As a result of the research, the following conclusions have been made:

– the most rational method for reducing greenhouse gas emissions at the current stage of technological development has been identified. The use of a mixture of methane and low-calorific gas—such as the product of wood waste gasification—helps to reduce the portion of carbon dioxide in flue gases that is related to greenhouse gases. Moreover, with the high cost of natural gas, replacing part of it with cheaper low-calorific gas can lead to a reduction in the cost of the produced energy;

– a method and model have been proposed for calculating the composition of flue gases with varying ratios of methane and low-calorific gases in their mixture. This method allows for determining both the characteristics of the combustible gas mixture based on the given parameters of the flue gases, as well as the amount of flue gases for specified fuel characteristics;

– two possible methods of controlling the power of the installed equipment by adjusting the share of low-calorific gas in the fuel mixture were considered. It was found that reducing power while attempting to maintain a constant volumetric flow rate of exhaust gases is not feasible. Moreover, in this scenario, the power (and the adiabatic combustion temperature) increases as the share of low-calorific gas rises. The reason for this effect was identified, highlighting the need for modernization of the fuel gas system and an increase in the heat exchange surface area (economizer);

– the possibility of controlling power while maintaining a constant volumetric flow rate of the combustible gas mixture was identified. As the share of low-calorific gas increases, the equipment's power decreases. It was demonstrated that it is possible to maintain exhaust gas parameters comparable to those observed when using pure methane. This feature enables the use of existing equipment without modernization for burning gas mixtures, while simultaneously reducing the share of carbon dioxide classified as a greenhouse gas.

References

1. Dash S.K., Chakraborty S., Elangovan D. A Brief Review of Hydrogen Production Methods and Their Challenges. *Energies*. 2023. V.16(3). 1141. 1-17. URL: https://doi.org/10.3390/en16031141
2. Al-Qahtani A., Parkinson B., Hellgardt K., Shah N., Guillen-Gosalbez G. Uncovering the true cost of hydrogen production routes using life cycle monetization. *Applied Energy*. 2021, V.281. 115958. 1-12. URL: https://doi.org/10.1016/j.apenergy.2020.115958
3. Guerra O.J., Eichman J., Kurtz J., Hodge B.-M. Cost Competitiveness of Electrolytic Hydrogen. *Joule.* 2019. V.3(10). P.2425-2443. URL: https://doi.org/10.1016/j.joule.2019.07.006
4. Pelykh S.N., Maksimov M.V. The method of fuel rearrangement control considering fuel element cladding damage and burnup. Problems of Atomic Science and Technology. 2013. 87(5). P.84-90. URL: https://vant.kipt.kharkov.ua/TABFRAME.html
5. Maksimov M.V., Pelykh S.N., Gontar R.L. Principles of controlling fuel-element cladding lifetime in variable VVER-1000 loading regimes. *Atomic Energy*. 2012, V.112(4). P. 241-249. URL: https://doi.org/10.1007/s10512-012-9552-3
6. Foshch T., Maksimov M., Pelykh S., Maksimova O. Models and methods for automated control of power change at VVER-1000 nuclear power unit. *Nuclear and Radiation Safety.* 2018. 1(77). P. 24-30. URL: https://www.researchgate.net/publication/326842689_
7. Foshch T., Machado J., Portela F., Maksimov M., Maksimova O. Comparison of two control programs of the VVER-1000 nuclear powerp.unit using regression data mining models. *Nuclear and Radiation Safety*. 2017. V.3(75). P. 11-17. URL:

https://www.sciencedirect.com/science/article/pii/S1877050916323195
8. Soroka B. S., Pyanykh K. Ye., Zgurskyi V. O. Mixed Fuel for Household Gas-Powered Appliances as an Option to Replace Natural Gas with Hydrogen. *Science in innovation*. 2022. V.18(3). P.10—22. URL: https://doi.org/10.15407/scine18.03.010
9. Xin Y., Wang K., Zhang Y., Zeng F., He X., Takyi S.A., Tontiwachwuthikul P. Numerical Simulation of Combustion of Natural Gas Mixed with Hydrogen in Gas Boilers. *Energies*. 2021.V.14. 21: 6883. URL: https://doi.org/10.3390/en14216883
10. Maksymov M., Lozhechnikov V., Maksymova O., Lysiuk O. Improvement of the control system over drum boilers for burning combustible artificial gases. *Eastern-European Journal of Enterprise Technologie*s. 2017. V.4(8-88). P.10-16. https://doi.org/10.15587/1729-4061.2017.107358
11. Taler J., Trojan M., Dzierwa P., Kaczmarski K., Węglowski B., Taler D., Jaremkiewicz M. The flexible boiler operation in a wide range of load changes with considering the strength and environmental restrictions. *Energy*. 2023. V.263, B, 125745. https://doi.org/10.1016/j.energy.2022.125745
12. Maya L., Restrepo A., Amell Arrieta A.A. Theoretical and numerical study of the combustion properties of premixed hydrogen/natural gas/air at a sub-atmospheric pressure of 0.849 Bar. *CT&F–Ciencia, Tecnología Y Futuro*. 2021. V.11(2). P. 39–49. URL: https://doi.org/10.29047/01225383.374
13. Fedorova N., Aziziyanesfahani P., Jovicic V., Zbogar-Rasic A., Khan M., Delgado A. Investigation of the Concepts to Increase the Dew Point Temperature for Thermal Energy Recovery from Flue Gas, Using Aspen. *Energies*. 2019. V.12. 1585. https://doi:10.3390/en12091585
14. Bălănescu D.T., Homutescu V.M. (). Effects of hydrogen-enriched methane combustion on latent heat recovery potential and environmental impact of condensing boilers. *Applied Thermal Engineering*. 2021. V.197, 117411. URL: https://doi.org/10.1016/j.applthermaleng.2021.117411
15. Dudynski M. Novel oxygen-steam gasification process for high quality gas from biomass. *Detritus*. 2018. 06.2019(0),1, P. 68-76. URL: https://doi.org/10.31025/2611-4135/2019.13814
16. Лавренцов Є., Сігал І., Сміхула А., Домбровська Е., Кернажицька О., Марасін А.
17. Досвід розробки, впровадження та модернізації водогрійних котлоагрегатів з двосвітними екранами та щілинними подовими пальниками. *Енерготехнології та ресурсозбереження.* 2019. №3. С. 17-26. URL: https://doi.org/10.33070/etars.3.2019.02
18. Сігал І.Я., Сміхула А.В., Марасін, А.В., Лавренцов Є.М., Домбровська Е.П. Модернізація газових котлів ТЕС, ТЕЦ та котелень відповідно до вимог екологічних директив ЄС. *Енерготехнології та ресурсозбереження.* 2017. Т.4. С. 61-71. http://nbuv.gov.ua/UJRN/ETRS_2017_4_12
19. Brunetkin O., Maksymov M., Maksymova O., Zosymchuk A. Development of the method of approximate solution to the nonstationary problem on heat transfer through a flat wall. *Eastern-European Journal of Enterprise Technologies*. 2017. V.6(5-90), P.31-40. https://doi.org/10.15587/1729-4061.2017.118930
20. Brunetkin O., Maksymov M., Maksymova O., Zosymchuk A. Development of a method for approximate solution of nonlinear ordinary differential equations using pendulum motion as an example. *Eastern-European Journal of Enterprise Technologies*. 2017. V.5(4-89). P. 4-11. URL: https://doi.org/10.15587/1729-4061.2017.109569
21. Brunetkin O., Maksymov M.V., Maksymenko A., Maksymov M.M. Development of the unified model for identification of composition of products from incineration, gasification, and slow pyrolysis. *Eastern-European Journal of Enterprise Technologies*. 2019. V.4(6-100). P. 25-31. URL: https://doi.org/10.15587/1729-4061.2019.176422
22. Bezhan V., Zhitarenko V. Modeling and Analysis of Energy Efficiency Parameters of Medium Pressure Boilers Using a Mixture of Natural and Blast Furnace Gases Taking into

Account Air Intakes. *Вісник Національного технічного університету «ХПІ». Серія: Енергетичні та теплотехнічні процеси й устаткування*. 2020. №2, С.32–39. URL: https://doi.org/10.20998/2078-774X.2020.02.05

МОДЕЛЬ МІНІМІЗАЦІЇ ВИКИДІВ ПАРНИКОВИХ ГАЗІВ ПРИ ЗМІННОМУ НАВАНТАЖЕННІ ПАРОВОГО КОТЛА

Т. Д. Марколенко, Д. Г. Проданов

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Email: tanyadmb@ukr.net

Досліджено проблему зниження викидів парникових газів під час змінного навантаження парогенеруючого енергоустаткування. Мета дослідження полягає у розробці методу та моделі, які дозволяють знизити викиди парникових газів шляхом використання суміші регульованого складу метану та альтернативного низькокалорійного газу. Наукова і практична значущість роботи полягає у створенні методів регулювання, які забезпечують екологічність і енергоефективність обладнання без його модернізації. Методологія дослідження базується на математичному моделюванні процесу горіння газової суміші із застосуванням рівнянь хімічної кінетики для ізоентальпійного процесу. У межах дослідження розроблено модель, яка дозволяє з високою точністю визначати адіабатичну температуру процесу горіння паливно-повітряної суміші, а також кількісний і якісний склад продуктів згоряння, що утворюються за умов використання різного співвідношення метану та низькокалорійного газу. Виявлено залежність між витратою паливної суміші, її складом і сталістю об’єму продуктів згоряння. Показано, що підтримання постійної витрати димових газів можливе шляхом регулювання складу паливної суміші, що впливає на потужність обладнання та температуру відхідних газів. Окремо розглянуто вплив дотримання умов ізоентальпійності процесу на зміну кількості виділеної енергії, що дозволило визначити закономірності регулювання потужності обладнання без значних втрат енергоефективності. Цінність дослідження полягає у розробці нових підходів до керування енергоефективністю та скорочення викидів парникових газів. Практичне значення полягає у можливості застосування отриманих результатів для регулювання потужності обладнання без його модернізації, зберігаючи сталий обсяг продуктів згоряння.

Ключові слова: регулювання потужності, низькокалорійний газ, парникові гази, хімічна кінетика, ізоентальпійний процес.

DOI 10.15276/imms.v14.no4.296
УДК 004.621

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 296-304

DEVELOPMENT OF COMPUTERIZED TECHNOLOGY FOR CREATING INDIVIDUAL RESPIRATORY PROTECTION EQUIPMENT USING 3D MODELING AND CAD

V. M. Tigariev, O. S. Lopakov, A. S. Koliada, V. V. Kosmachevskiy

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Emails: tigarev.v.m@op.edu.ua, lopakov.o.s@op.edu.ua,
a.s.koliada@op.edu.ua, kosmachevsky.v.v@op.edu.ua

The development of individual respiratory protective equipment (masks) during the coronavirus pandemic is a pressing issue. Modern design and manufacturing technologies make it possible to create masks that account for the unique anatomical features of each individual. However, existing mask configurations often fail to consider the specific characteristics of the wearer. Therefore, it is critically important to design masks tailored to the anatomical features of each person. This paper examines the design process for protective masks using modern computer technologies and an information-based model. The general approach and practical options for mask design are presented, taking into account the individual's features. The information model comprises six main stages for creating a protective face mask. To design a human head form, the photogrammetry method is employed, allowing the creation of a three-dimensional head form from two-dimensional photographs. Using retopology tools on the surface of the 3D head form, the basis for the mask frame is developed in the 3DS Max program. Subsequently, a three-dimensional solid model of the mask frame is created. The process of generating the solid model and testing the mask frame under mechanical stresses, such as changes in facial expressions, is conducted in Autodesk Inventor. To enhance the mask's secure attachment to the face, a version with ear hooks, similar to those used in eyeglasses, is proposed. Additionally, the design allows for the use of one or two filters, as in the Pitta mask. An automation subsystem for mask design is developed in the iLogic environment of Autodesk Inventor, based on the results of this study.

Keywords: Information model, mask frame, filtering element.

Introduction. During the coronavirus pandemic, the creation of customized respiratory protective equipment (face masks) has become a highly relevant issue. Modern computer-aided design and manufacturing technologies enable the production of masks tailored to the individual anatomical features of each person.

Designing face masks adapted to individual needs is not only important during pandemics but also under normal circumstances, such as for surgeons during surgeries, police officers while on duty, and other professionals requiring specialized protective equipment [1].

Literature Review. All individual protective products are divided into two major groups [2, 3]: disposable and reusable. Products from the first group are intended for single use and are disposed of afterward, while products from the second group are designed for long-term use, lasting up to several months. In this regard, reusable masks typically have lower penetration rates but come at a higher price. Both types of respiratory protective equipment are used not only for personal purposes but also in healthcare institutions, beauty salons, factories, and in cases of man-made threats. There is an extensive classification of these products.

Disposable masks are commonly used in the medical sector, cosmetology, and daily life. The maximum recommended duration for wearing such a mask is no more than 2 hours, after which it should be replaced with a new one. Traditionally, these masks consist of an outer layer and a filtering layer. Additionally, they may include a hydrophobic layer or film to prevent

glasses from fogging. Some masks are equipped with a flexible aluminum strip to ensure a tight fit around the nose.

By filtration level, disposable masks are classified as follows:

- Double-layered: Basic masks with a protection level of up to 98%;
- Three-layered: Designed for everyday use, with a filter placed at the center;
- Four-layered: Surgical masks that provide protection against liquid penetration.

By material, masks are produced using:

- Cotton: Utilizing cotton filters;
- Spunbond: Offering high air permeability;
- Meltblown: Featuring an inner filtering layer;
- Self-made: Made from materials such as gauze, cotton, or linen.

By availability of additional features:

- With a valve: Includes an adjustable opening for moisture removal;
- Without a valve: Fabric with heat-absorbing properties;

In addition to these categories, reusable masks offer greater comfort and ease of maintenance [4]. They can be washed at high temperatures after each use without losing their protective effectiveness. For instance, the so-called Pitta mask is a filter made of foamed polyurethane that fits snugly to the face. This type of reusable mask can retain up to 90% of pathogenic microparticles. It can be washed and reused after drying but has a maximum lifespan of two months. These masks are significantly more expensive than disposable ones.

Masks with filters can provide up to five levels of protection:

Level 1 (air-permeable forming layer): Blocks larger particles;

Level 2 (activated carbon): Adsorbs chemical and virological pollutants;

Level 3 (powdered cotton): Filters out smaller particles;

Level 4: Enhances particle filtration;

Level 5 (nonwoven breathable material): Ensures breathability while maintaining filtration.

Contests for designing innovative types of masks are held in various countries. Experts assert that textile masks do not provide reliable protection against coronavirus. More effective protective equipment against biological threats includes respiratory masks, specialized filters, protective screens, and masks with advanced coatings. In the first year of the pandemic, numerous innovative protective masks were developed worldwide [5].

Papers [1-4] have reviewed different types of masks and provided recommendations based on expert opinions. However, these studies do not consider the individual anatomical features of the human face. The approach proposed in this study aligns with the second phase of competition: the creation of a new concept for protective masks.

Research Methodology. Nowadays, creating new objects in various fields, including medicine, is increasingly carried out using computer design technologies. The suggested approach for creating a mask involves the fulfillment of several successive stages. In the mask design and production process, various computer programs are applied. When designing new objects using computer technologies, it is important to build an information model. The information model generalizes the approach for design by applying different technologies, examples of which are presented in studies [7-9]. To optimize the design process, it is suggested to use an information model for creating face masks, which will contain all the information necessary for mask production. The result is developed into a mask design subsystem that considers the individual features of the human face. The design information model consists of six main stages, each divided into several sequential steps.

Let us describe in detail an algorithm of the information model for the discussed issue:

1. Stage – Data gathering and determining technology for creating a model of a human face.
    a. Determining the technology for producing a three-dimensional model of a human face.

   b. Determining equipment for scanning (photogrammetry) of the object under investigation.
   c. Conducting photogrammetry of the object and obtaining a mathematical model of the information model boundaries in the form of a point cloud.
2. Stage – Creating and optimizing the face model.
   a. Creating the initial three-dimensional surface model of the human head.
   b. Correcting form inconsistencies and optimizing the surface model of the face.
3. Stage – Modeling the mask configuration.
   a. Determining the technology for designing the mask frame.
   b. Retopology of the mask frame on the face model.
   c. Developing the model of the mask frame.
4. Stage – Form analysis and conducting stress simulations of the mask configuration.
   a. Analyzing the form and required stresses for the created solid mask model.
   b. Conducting static and dynamic stress simulations on the solid mask model.
5. Stage – Preparing mask production technology.
   a. Designing the mask configuration, taking into account the shape of the filtering element.
   b. Preparing the mask production technology based on the developed model.
   c. Creating the mask according to the developed technology.
6. Stage – Implementing and developing the automation subsystem for designing masks with different filter types.
   a. Summarizing versions of the mask configuration with different filter types.
   b. Conducting CAD analysis with an option for developing automation subsystems for object design.
   c. Developing an automation subsystem for designing masks with various frame configurations and filter types.

Let us consider in detail the general approach for mask design using the suggested information model.

At the first stage of creating individually tailored masks, we must obtain a 3D headform of a human face for which the mask will be designed. At this point, we are discussing building a 3D model, the real-life analogue of which cannot have an accurate description due to the infinite number of unique parameters. For such cases, 3D scanning and photogrammetry methods have been developed. The 3D scanning method requires specialized and often expensive equipment, such as 3D scanners. Photogrammetry, on the other hand, can be applied using consumer smartphones. It tracks points from different sources under consistent conditions and generates a 3D point cloud that forms the topology of the object. Information from each photograph is saved in a file, including height, camera rotation angle, and geospatial data. The program applies computer vision and photogrammetry techniques to find common points in multiple photographs. As a result, each pixel in a photograph is matched by color correspondence with others, and every match becomes a key point. If the key point is found in three or more photographs, the program constructs this point in space. The more such points, the better the coordinates of each point are defined in space. Therefore, the more overlap among photographs, the more accurate the resulting model. An overlap of 60–80% is considered optimal. The spatial coordinates of each point are calculated using the triangulation method: from every shooting point, a line of sight is automatically drawn to the selected point, and their intersection determines the required outcome [6]. Additionally, algorithms aimed at minimizing the sum of squares of errors are applied in photogrammetry. The Levenberg-Marquardt algorithm (or the damped least-squares method) is typically used, solving nonlinear equations by the least squares method. During the processing of photographs, an extended point cloud (a set of all 3D points) is created, which can be used to generate a surface consisting of polygons. Finally, resolution is calculated to determine which pixels in a photograph correspond to which

polygons. To do this, the 3D model is unfolded into a surface, and the spatial position of a point is matched to the original photograph to set its color.

For example, the object is a human face. To make this method work, the face should be photographed from different angles under consistent conditions: the same state of the object (facial expression, position), the same surroundings, and the same material. Maintaining these conditions guarantees a more successful search for the same points in different photographs, enabling a more accurate representation of the human face in a 3D point cloud [10].

At the second stage, photogrammetry software is used to create a 3D model of the face. The obtained model is then imported into software designed for processing three-dimensional models, such as Autodesk 3DS Max or Maya. At this point, form inconsistencies are corrected, and the surface model of the object is optimized, ensuring the correct form of the created object. At the third stage, the mask configuration is modeled based on the obtained face model. This involves solving an inverse problem: the geometry of the mask structure is placed over the face. Retopology tools [11-12] can be used for this purpose. Retopology is the process of creating a new geometry over an existing one by altering its structure. First, a surface tangential to the face is built. The created surface of the mask base is then imported into a CAD system, where a three-dimensional solid model of the mask frame is developed.

At the fourth stage, form analysis and stress simulations of the mask configuration are conducted. The material type and its properties are selected for the solid model. Modern CAD systems allow simulations of various stress types on created models. Stress analysis is required to account for facial expression changes during conversation, ensuring the mask fits tightly to the face.

At the fifth stage, mask production technology is prepared. Appropriate types of filtering elements are selected. It is suggested to use a replaceable filtering element that can be temporarily fixed to the basic mask frame using a sticky layer or a fixing mechanism. This allows the filtering element to be replaced after two hours of use. Alternatively, separate filters, such as those used in Pitta masks, can be applied, which can be washed in disinfectant fluid or replaced as needed. Once the filtering element type is defined, the mask production technology must be determined. Additive manufacturing using 3D printers is considered one of the most appropriate solutions. Modern 3D printers can use various materials to form an object [13, 14]. After producing the mask frame, an elastic layer is applied to the inner surface to reduce skin irritation and discomfort during extended use.

At the sixth and final stage, an analysis of available mask frame configurations and filtering elements is performed. The most suitable CAD system is selected for developing an automation design subsystem for the product.

Therefore, the general approach for creating protective masks using an information model has been studied in detail.

Results. The design of the protective mask using an information model is demonstrated in the presented example.

At the first stage, the required software for implementing photogrammetry is identified. To create a three-dimensional headform using the photogrammetry method, a variety of programs can be used.

In this research, the construction is performed using Regard3D, a free, open-source photogrammetry program. To create the headform, 100 photographs were taken, based on which the model boundaries are defined, and a point cloud of the future object is formed. In Regard3D, the point cloud is generated and serves as the core for the headform design.

At the second stage, the point cloud is edited in the Regard3D program before being forwarded to generate the three-dimensional surface of the object.

As a result of photogrammetry and point cloud correction, a three-dimensional surface headform is obtained (Fig. 1). Further editing of the headform is carried out using another software product – Autodesk 3DS Max.

Fig. 1. Three-dimensional surface model of a human face

At the third stage, the generated headform is imported into Autodesk 3DS Max, where the mask frame design process begins. For this purpose, it is recommended to use retopology tools. Retopology enables the creation of new geometry over an existing one by modifying its structure.

Using retopology tools, new geometry is placed over the face model, using it as a foundation (Fig. 2). This approach allows for precise placement of the new object and facilitates further modifications. It is also crucial to maintain the correct topology of the 3D model, which should consist exclusively of quadrilaterals. This ensures compatibility with subdivision modifiers, enabling the creation of a high-poly model while allowing a return to its previous state for further updates, if needed.

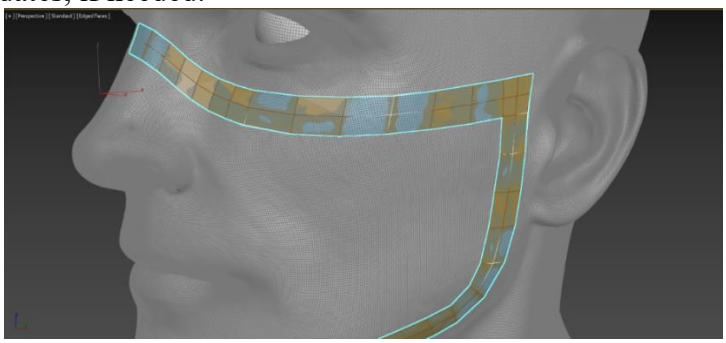


Fig. 2. Mask frame retopology

This toolbox facilitates the design of the mask frame shape, which can be easily modified later. Using the obtained mask frame configuration, a solid model is developed in Autodesk 3DS Max (Fig. 3). Managing the model's topology allows the design to be exported into CAD modeling programs (experiments in this study were conducted using the Autodesk Inventor Pro CAD system). This enables the performance of various stress tests, the selection of physical materials, the configuration of settings for 3D printing, and more.

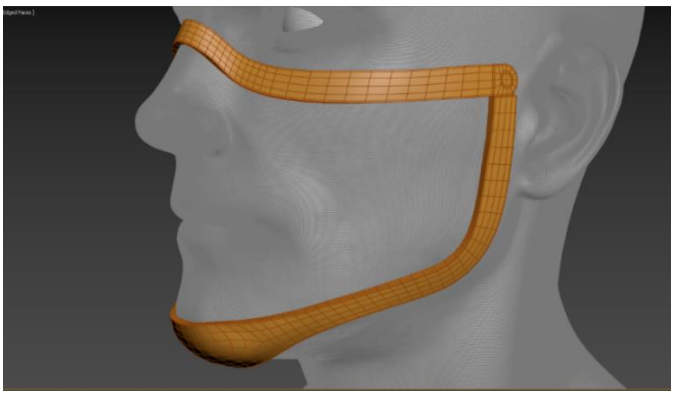


Fig. 3. Solid model of a mask frame in Autodesk 3Ds Max

At the fourth stage, the reliability of the mask structure is evaluated. The configuration of the mask frame consists of two main parts, which are hingedly connected. This design ensures that the mask can be used regardless of changes in facial expression or speech by the wearer. The lower part of the mask fits snugly against the chin, while the upper part is secured on the bridge of the nose.

Stress simulations are performed on all parts of the mask and their connections using Autodesk Inventor Pro software. These simulations verify the reliability of the mask configuration under mechanical stress and evaluate its fixation on the face. Once the material for the mask's basic structure is determined, the design is tested for resistance to external forces and potential displacement caused by changes in facial expressions during conversation. Adjustments to the mask structure can be made based on the testing results.

A three-dimensional solid model of the mask frame created in Autodesk Inventor Pro is shown in Fig. 4.

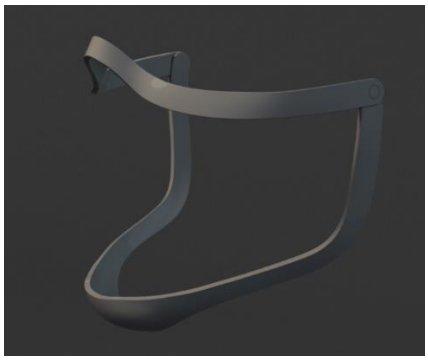


Fig. 4. Three-dimensional solid model of a mask frame in Autodesk Inventor Pro

At the fifth stage, the type of filtering element is defined, which determines the method of attaching it to the basic mask structure. Let us consider the simplest example involving a replaceable multi-layered filtering element. This element can be pre-manufactured, taking into account the shape of the basic mask configuration.

The filtering element is attached to the basic structure using a sticky layer, which can be applied to either the filtering element, the basic structure, or both. To ensure more reliable fixation, an additional securing element can be applied over the filter.

If necessary, the mask configuration can be adjusted to improve the secure attachment of the filtering element. Separate filters, similar to those in Pitta masks with one or two filters, can also be used.

Once the filtering element is defined, the technology for producing the mask frame is developed. Additive manufacturing using a 3D printer is selected as the core technology in this experiment.

The process of creating the object using a 3D printer can be simulated in the Autodesk Inventor HSM program. Using cloud computing technologies, the parameters of the required

3D printer are downloaded, and the simulation of the mask frame manufacturing process is conducted.

If complications arise during the simulation, the program suggests appropriate configuration corrections or the addition of extra fixing elements. The final version of the mask after the simulation of 3D printing is shown in Fig. 5.

Once the mask is designed, it is placed on a headform for testing. During the tests, the mask is carefully positioned to ensure accurate placement. The resulting image of the mask on a face is shown in Fig. 6.

To guarantee more secure fixation, a version of the mask with ear hooks, similar to those used in eyeglasses, is proposed (Fig. 7). Variants with separate filtering elements are presented in Fig. 8: (a) single filter and (b) double filter.

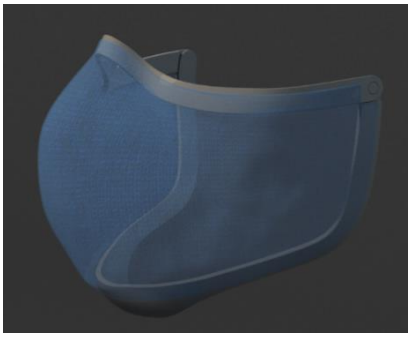


Fig. 5. Finished mask

Fig. 6. Mask put on a face

Fig.7. Mask with ear hooks

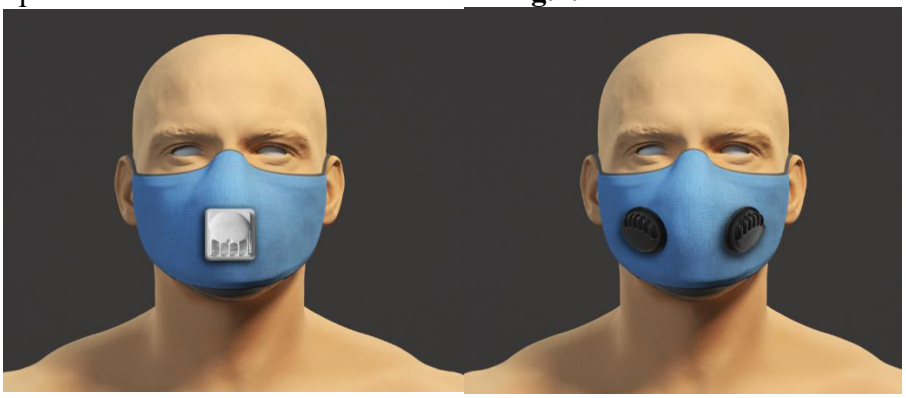


Fig. 8. a) Single filter mask b) Double filter mask

At the final sixth stage of the information model, an automation subsystem for designing masks is developed based on the required parameters obtained during the previous stages. The integrated design automation tool in the Autodesk Inventor CAD system, the iLogic environment, was used to create this subsystem. The presented subsystem streamlines the process of generating an appropriate template for the face mask. An interface window of this subsystem is shown in Fig. 9

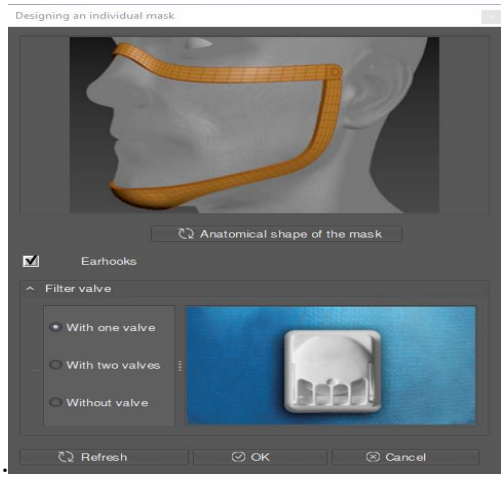


Fig. 9. Subsystem interface window

Conclusions. This study presents a general approach to protective mask design that considers the individual features of the human face. An information model of the created object is utilized in the mask design process. The information model encompasses all the sequential stages of mask creation and includes the possibility of developing an automated design subsystem. The construction of real protective mask prototypes with different replaceable filters is discussed in detail. To improve the reliability of mask fixation on the face, the addition of ear hooks, similar to those used in eyeglasses, is suggested.

Applying an information model in the mask creation process reduces the time required for design, increases accuracy, and helps avoid errors in the developed model. The presented versions of protective face masks address the issues outlined in the Mask Innovation Challenge program. The innovative mask design method based on the information model can also be applied to other similar objects. The development of the mask design subsystem will facilitate their implementation. Further research on this topic aims to design reusable masks that can accommodate a variety of filter types.

References

1. Coronavirus Face Masks: What You Should Know. URL: https://www.webmd.com/lung/coronavirus-face-masks#6
2. Best face masks to wear: Doctors share their favorite masks URL: https://www.nbcnews.com/shopping/apparel/best-face-masks-doctors-n1257689
3. N95 Respirators, Surgical & Face Masks URL: https://www.fda.gov/medical-devices/personal-protective-equipment-infection-control/n95-respirators-surgical-masks-and-face-masks
4. Behind the mask: Rethinking the use of face masks while exercising. N. Lakicevic, G. D'Antona, A Paoli, A Bianco, N Maksimovic, S Ostojic, P Drid. URL: https://www.researchgate.net/publication/350959198_Behind_the_mask_Rethinking_the_use_of_face_masks_while_exercising
5. Challenge. URL: https://www.challenge.gov/
6. R. Larsson Methodology for Topology and Shape Optimization: Application to a Rear Lower Control Arm Geteborg, Sweden: Chalmers University of Technology, 2016. URL: https://odr.chalmers.se/server/api/core/bitstreams/479ca2f9-b3d0-42d0-ab95-b4a38fc0e9ca/content
7. Tigariev V., Barchanova Y., Prokopovych I., Lopakov O., Vinokurov R. An individual mask creation using the information model. *Proceedings of Odessa Polytechnic University*. 2021. No 1(63). P. 95–105. URL: https://pratsi.op.edu.ua/app/webroot/articles/1632908067.pdf

8. Tigariev V., Salii V., Rybak O., Barchanova Y., Lopakov O. Reverse Engineering Based on Information Model. *Lecture Notes in Mechanical Engineering*. 2021. P. 217–226 DOI:10.1007/978-3-030-68014-5_22
9. Wagner G. Information and Process Modeling for Simulation. *Journal of Simulation Engineering*. 2018. V.1(1) P. 1–25. URL: https://jsime.org/index.php/jsimeng/article/view/2
10. Regard3d. URL: http://www.regard3d.org/index.php
11. Head Modeling with 3dsmax. URL: https://3dtotal.com/tutorials/t/head-modeling-with-3dsmax-hatice-bayramoglu-character-human-face
12. Retopology of human face in 3d coat, texture clean up in photoshop. URL: https://www.cgtrader.com/tutorials/3843-retopology-of-human-face-in-3d-coat-texture-clean-up-in-photoshop
13. Kamonichkin D.T. 3D printing of surgical templates. 2018.. URL: https://studia3d.com/hirurgshablon/
14. Sergeev A.N. 3D printers in medicine - a field of application and prospects for the development of printing . 2017. URL: https://www.sciencedebatecom/3d-printery-v-meditsine-oblast-primeneniya-i-perspektivy-razvitiya-pechati/

РОЗРОБКА КОМП'ЮТЕРИЗОВАНОЇ ТЕХНОЛОГІЇ СТВОРЕННЯ ІНДИВІДУАЛЬНИХ ЗАСОБІВ ЗАХИСТУ ОРГАНІВ ДИХАННЯ З ВИКОРИСТАННЯМ 3D МОДЕЛЮВАННЯ ТА САПР

В. М. Тігарєв, О. С. Лопаков, А. С. Коляда, В. В. Космачевський

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Emails: tigarev.v.m@op.edu.ua, lopakov.o.s@op.edu.ua, kosmachevsky.v.v@op.edu.ua

Розробка індивідуальних засобів захисту дихання (маски) під час пандемії коронавірусу дуже актуальна. Сучасні технології проектування та виготовлення дозволяють створити маски з урахуванням індивідуальних анатомічних особливостей кожної людини. Докладно розглянуті різні типи захисних масок та їх конструкції. Наявні конструкції масок не враховують індивідуальних особливостей кожної людини. Тому надзвичайно важливо створити маску відповідно до анатомічних особливостей кожної людини. У даній роботі розглядається процес створення захисної маски із використанням сучасних комп'ютерних технологій із застосуванням інформаційної моделі. Розглянуто загальний підхід та варіанти практичної реалізації створення маски з урахуванням індивідуальних особливостей людини. Інформаційна модель включає шість основних етапів створення захисної маски. Для створення моделі голови людини використано метод фотограмметрії, коли за двовимірними фотографіями формується тривимірна модель голови людини. За допомогою технології ретопології на поверхні тривимірної моделі голови у програмі 3DS Max створюється основа несучого каркаса маски. Створення твердотільної моделі та перевірка несучого каркаса маски на механічні навантаження при зміні міміки обличчя проведено в Autodesk Inventor. Для більш надійного кріплення маски на обличчі запропоновано варіант із завушинами, як у окулярів. Розглянуто варіант встановлення одного або двох фільтрів як у масці Піття. За результатами дослідження було створено підпрограму автоматизації проектування створеної маски у середовищі iLogic Autodesk Inventor.
Ключові слова: інформаційна модель, каркас маски, фільтрувальний елемент.

DOI 10.15276/imms.v14.no4.305
УДК 004.056.55

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 305-313

A METHOD FOR POLYNOMIAL RECOVERY FROM ITS RESIDUES BASED ON ADDITION IN Z[x] RING

I. Yakymenko, M. Kasianchuk, I. Shylinska

West Ukrainian National University
11, Lvivska str., Ternopil, 46009, Ukraine
Emails: iyakymenko@ukr.net, kasyanchuk@ukr.net

The methods for polynomial recovery from its residues in Z[x] ring are presented in this paper. This problem is relevant due to important applications in asymmetric and symmetric cryptography, algorithms of noise resistant coding, in the process of transmitting data packets for error control and recovery in computer networks and distributed data storage. The theoretical foundations of polynomial recovery in a ring of polynomials based on known approaches, namely, on the Chinese Remainder Theorem and Garner's algorithm, are considered, and their advantages and disadvantages are highlighted. New methods of inverse transform from the Residue Number System based on addition of the product of moduli and the product of polynomial residues are developed. Analytical expressions of the time complexity of the proposed method and Garner's algorithm are created. The graphs of their dependences are presented, which show that the developed method for polynomial recovery from its residues in Z[x] ring is characterized by lower complexity. It was found that the time complexities of both methods increase with an increase in the dimensions of the input parameters. The efficiency of the use of the developed method in the ring Z[x] is studied, which shows a logarithmic growth with an increase in the degrees of the polynomial, and a proportional decrease when a number of moduli increases.
Key words: polynomial recovery, ring of polynomials, residues, Chinese Remainder Theorem, Garner's algorithm, time complexity, efficiency.

Introduction. Polynomial recovery from its residues in a ring of polynomials is an important problem of modern algebra and number theory [1,2]. In practice, similar to the case of integers [3-5], the application of this theory, namely the Residue Number System (RNS) in a polynomial ring (PR), allows working not with higher degree polynomials, but with sets of residues whose degree is less than or equal to the selected moduli (irreducible polynomials) [6, 7]. One of the main advantages of using RNS PR is that calculations can be performed in parallel for each module [8]. These properties make it possible to reduce the complexity of calculations and, accordingly, to increase the efficiency of computer systems due to the parallel process of performing arithmetic operations [9], to control and correct errors in noise resistant coding [10, 11], and etc. Polynomial Residue Number Systems (PRNS) are widely used in modern cryptography, in particular, in RSA [12, 13], Rabin [14], AES ciphers [15], and etc. In the RSA cryptosystem, residues on division by a high-order polynomial, which is the product of two irreducible polynomials, are calculated [13]. Accordingly, to increase performance, calculations can be carried out modulo irreducible polynomials of a significantly lower order. Therefore, the development of methods and algorithms that make it possible to reduce the time complexity when recovering a polynomial from its residues in a polynomial ring is currently a relevant problem.

Related Works. The main ideas of the PRNS were outlined in [9, 16]. The most important contribution of these works is the fundamental theorem on the modular number system. Due to the theorem, the coefficients of the polynomials used to represent the set Zp, are restricted. In [17], the requirements for the size of the parameters used to represent an integer modulo p were defined. Work [18] is devoted to the development of multiplication algorithms in PRNS. In [19], a modified Chinese Remainder Theorem (CRT) for cyclotomic polynomials was presented, which made it possible to simplify the polynomial recovery from its residues. After

that, PRNS found its application in problems of noise resistant coding. In particular, in [20], the principles of generating redundant codes in the polynomial number system were considered, and the algorithm was developed, which made it possible to detect and correct errors without the inverse transform of the PRNS code into a positional one and without performing a division operation. It should be noted that the Chebyshev polynomials take an important role in data protection. For example, in [21], a scheme in the client-server environment was proposed on the basis of the Chebyshev polynomials. Although this scheme demonstrates a lower speed compared to the existing ones, it can resist some popular attacks. Recently, the Chebyshev polynomials have been actively used in asymmetric cryptography. Thus, in [22], two asymmetric cryptosystems were developed on the basis of the Chebyshev polynomials, the main feature of which is the generation of a semigroup with respect to the composition operation. An image encryption algorithm based on the Chebyshev polynomials was proposed in [23]. In [24], modified fast algorithms of asymmetric cryptography, i.e., matrix algorithm and algorithm based on the characteristic polynomial were developed and an efficient scheme for calculating the Chebyshev polynomials over a finite field was presented.

Methods for polynomial recovery. The theoretical basis for polynomial recovery from its residues in the corresponding ring, as for integer arithmetic [25], is algebra and number theory, in particular the CRT in a polynomial form. Any polynomial $N(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ can be represented in the form of residues $b_i(x) = c_l x^l + c_{l-1}x^{l-1} + \ldots c_1 x + c_0$ of division by irreducible polynomials $p_i(x) = s_z x^z + s_{z-1}x^{z-1} + \ldots + s_1 x + s_0$, which are called polynomial moduli

$$b_i(x) = N(x) \bmod p_i(x), \tag{1}$$

where $\deg b_i = \deg N - \deg p_i$.

At the same time, a necessary and sufficient condition is the inequality $N(x) < P(x) = \prod_{i=1}^{k} p_i(x)$, where k is the number of irreducible polynomials. Then, the original polynomial can be unambiguously recovered on the basis of the CRT:

$$N(x) = \left(\sum_{i=1}^{k} m_i(x) P_i(x) b_i(x) \right) \bmod P(x), \tag{2}$$

where $P_i(x) = \dfrac{P(x)}{p_i(x)}$, $m_i(x) = P_i^{-1}(x) \bmod p_i(x)$.

Another method of a polynomial recovery from its residues in the corresponding ring is Garner's algorithm, which is based on the relation:

$$N(x) = n_0(x) + n_1(x)p_1(x) + n_2(x)p_1(x)p_2(x) + \ldots + n_{n-1}(x)p_1(x)p_2(x)\ldots p_{i-1}(x), \tag{3}$$

where $0 \le n_i(x) < p_{i+1}(x)$, i=0, 1, …, k-1,

$$n_i(x) = \frac{b_{i+1}(x) - \left(n_0(x) + n_1(x)p_1(x) + \ldots + n_{i-1}(x)p_1(x)p_2(x)\ldots p_{i-1}(x)\right)}{p_1(x)p_2(x)\ldots p_i(x)} \bmod p_{i+1}(x). \tag{4}$$

In this case, the polynomials $n_i(x)$ are calculated sequentially one after another based on the recurrence formula (4). In addition, both Garner's algorithm and CRT can be used for similar operations in integer arithmetic.

The main disadvantage of the above methods of a polynomial recovery from its residues is strictly sequential structure of polynomials, which makes it impossible to parallelize calculations, perform operations on polynomials of higher orders (in particular, calculate the residue modulo $P(x)$), and it is necessary to find modular multiplicative inverse in the ring of polynomials, which is a cumbersome task even for integer arithmetic [26]. To find it, the following methods are most commonly used [27]: sorting through all possible options, using the extended Euclidean algorithm, based on the Euler function. These approaches are characterized by significant computational complexity.

Therefore, the purpose of this work is to develop the methods for polynomial recovery from its residues in a ring of polynomials based on the product addition and the addition of moduli residues with the possibility of parallelizing calculations and avoiding the multiplicative

inverse polynomial search procedure. At the same time, the results of intermediate calculations will not go beyond the set range, which eliminates the need to perform the operation of finding residue relatively large modulo $P(x)$.

Polynomial recovery method based on addition of the product of the polynomial moduli. Let us consider the system of congruences, which is built based on relation (1):

$$\begin{cases} b_1(x) \equiv N(x) \bmod p_1(x) \\ b_2(x) \equiv N(x) \bmod p_2(x) \\ \ldots\ldots\ldots\ldots\ldots \\ b_k(x) \equiv N(x) \bmod p_k(x). \end{cases} \quad (5)$$

Any congruence modulo $p_1(x)$ with the residue $b_1(x)$ (e.g., $a(x) \bmod p_1(x) \equiv b_1(x)$) can be represented in the form of $a(x) = \gamma(x) p_1(x) + b_1(x)$, where $\gamma(x)$ is a polynomial that indicates how many times modulo $p_1(x)$ must be added to the residue $b_1(x) = N_1(x)$ to satisfy the relation $N_2(x) \bmod p_2(x) \equiv b_2(x)$. In this case $N_2(x) = N_1(x) + \gamma_1(x) p_1(x)$. Next, it is necessary to add the product $p_1(x) p_2(x)$ until the congruence $N_3(x) \bmod p_3(x) \equiv b_3(x)$, where $N_3(x) = N_2(x) + \gamma_2(x) p_1(x) p_2(x)$, is fulfilled. This procedure continues until the last equation (5) is satisfied. Analytically, it is written as follows:

$$\begin{aligned} & N_1(x) = b_1(x); \\ & N_2(x) = N_1(x) + \gamma_1(x) p_1(x) = b_1(x) + \gamma_1(x) p_1(x);\ N_2(x) \bmod p_2(x) \equiv b_2(x); \\ & N_3(x) = N_2(x) + \gamma_2(x) p_1(x) p_2(x) = b_1(x) + \gamma_1(x) p_1(x) + \gamma_2(x) p_1(x) p_2(x);\ N_3(x) \bmod p_3(x) \equiv b_3(x); \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & N_i(x) = N_{i-1}(x) + \gamma_{i-1}(x) p_1(x) p_2(x) \ldots p_{i-1}(x);\ N_i(x) \bmod p_i(x) \equiv b_i(x); \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & N_k(x) = N(x) = N_{k-1}(x) + \gamma_{k-1}(x) p_1(x) p_2(x) \ldots p_{k-1}(x);\ N_k(x) \bmod p_k(x) \equiv r_k(x). \end{aligned} \quad (6)$$

The search for $\gamma_i(x)$ is carried out using the method of undetermined coefficients, and at each step of the algorithm, the degree of the polynomial $\gamma_i(x) = A_i x^i + A_{i-1} x^{i-1} + \ldots + A_1 x + A_0$ will be 1 degree less then $p_{i+1}(x)$: $\deg \gamma_i(x) = \deg p_{i+1}(x) - 1$.

Let us consider the following example. Let the given system of comparisons be:

$$\begin{cases} N(x) \bmod (x^2 + x + 1) \equiv x + 3 \\ N(x) \bmod (x^2 + x + 2) \equiv 2x + 5 \\ N(x) \bmod (x^3 + 2x + 1) \equiv x^2 + 4x + 1 \end{cases} \quad (7)$$

Without reducing the generality of the problem, we can assume that for polynomials $p_1(x), p_2(x), \ldots, p_k(x)$ their degrees satisfy the inequalities $\deg p_1 \geq \deg p_2 \geq \deg p_3 \geq \ldots \geq \deg p_k$. These conditions allow you to take larger steps when performing iterations.

Let us find the value of $a(x)$ in the relation $a(x) \bmod p_1(x) \equiv b_1(x)$. It can be presented in the form of $(\gamma(x)(x^2 + x + 1) + x + 3) \bmod (x^2 + x + 2) = 2x + 5$. Here, $\gamma(x)$ is a polynomial that indicates how many times modulo $p_1(x) = x^2 + x + 1$ must be added to the residue $x + 3 = N_1(x)$ in order to satisfy the relation $N_2(x) \bmod (x^2 + x + 2) \equiv 2x + 5$, where $N_2(x) = x + 3 + \gamma_1(x)(x^2 + x + 1)$. Since $p_1(x) = x^2 + x + 1$ is a polynomial of the second degree, it is advisable to present the found parameter $\gamma_1(x)$ in the form of a polynomial of the first degree: $\gamma_1(x) = A_1 x + A_0$. Then the product $\gamma_1(x)(x^2 + x + 1) = (A_1 x + A_0)(x^2 + x + 1) = A_1 x^3 + (A_1 + A_0) x^2 + (A_1 + A_0) x + A_0$ is obtained. To find the unknown coefficients A_i, consider the congruence $(A_1 x^3 + (A_1 + A_0) x^2 + (A_1 + A_0) x + A_0) \bmod (x^2 + x + 2) = ((x^2 + x + 2)(A_1 x + A_0) + (-A_1 x - A_0)) \bmod (x^2 + x + 2) = x + 2$. Therefore, taking into account the last expression, coefficients A_1 and A_0 take the values -1, -2, and, respectively, $\gamma(x) = -x - 2$, $N_2(x) = -x^3 - 3x^2 - 3x - 2$.

Next, it is necessary to add the product $p_1(x)p_2(x)=(x^2+x+1)(x^2+x+2)=(x^4+2x^3+4x^2+3x+2)$ until the congruence $N_3(x)\bmod p_3(x)=b_3(x)$ is satisfied, where $x^2+4x+1=((-x^3-3x^2-3x-2)+\gamma_2(x)(x^4+2x^3+4x^2+3x+2))\bmod(x^3+2x+1)$. First, the value of $(-x^3-3x^2-3x-2)\bmod(x^3+2x+1)=-3x^2-x-1$, which is placed on the left side of the latter equation, is calculated: $4x^2+5x+2=\gamma_2(x)(x^4+2x^3+4x^2+3x+2)\bmod(x^3+2x+1)$. Having reduced the modulo degree by 1, parameter $\gamma_2(x)$ should be searched for in the form of $\gamma_2(x)=A_2x^2+A_1x+A_0$, i.e., $4x^2+5x+2=(A_2x^2+A_1x+A_0)(x^4+2x^3+4x^2+3x+2)\bmod(x^3+2x+1)$. As $(x^4+2x^3+4x^2+3x+2)\bmod(x^3+2x+1)=2x^2-2x$, then $4x^2+5x+2=(A_2x^2+A_1x+A_0)(2x^2-2x)\bmod(x^3+2x+1)$ can be obtained. Next the value of $(A_2x^2+A_1x+A_0)(2x^2-2x)=2A_2x^4+2A_1x^3+(2A_0-2A_2)x^2-2A_1x-2A_0$ is calculated and the residue $(2A_2x^4+2A_1x^3+(2A_0-2A_2)x^2-2A_1x-2A)\bmod(x^3+2x+1)=(2A_0-6A_2)x^2+(-6A_1-2A_1)x-2A_0-2A_1$ is found. The search for unknown coefficients is reduced to the solution of the system of equations:

$$\begin{cases}2A_0-6A_2=4\\-6A_1-2A_2=5\\-2A_0-2A_1=2\end{cases},\ \text{or}\ \begin{cases}A_0-3A_2=2\\-6A_1-2A_2=5\\-A_0-A_1=1\end{cases}.$$

Its solutions are the values of $A_2=-\frac{13}{16}, A_1=-\frac{9}{16}, A_0=-\frac{7}{16}$. So, by substituting A_2, A_1, A_0, we get the recovered polynomial from its residues according to the proposed algorithm:

$$N(x)=(-x^3-3x^2-3x-2)+(4x^2+5x+2)(x^4+2x^3+4x^2+3x+2)=4x^6+13x^5+28x^4+35x^3+28x^2+13x+2$$

Therefore, the solution to system (7) is a polynomial, which is obtained without the use of cumbersome operations and control over the overflow of the bit grid when performing intermediate calculations.

It should be noted that the proposed method is similar to Garner's algorithm, but the operation of finding the modular inverse in the ring of polynomials to obtain the corresponding coefficients is eliminated.

The method for decimal number recovery based on addition of the residue from the product of moduli. To simplify the calculations used in the proposed method, it is possible to add not the product of polynomials- moduli, but the residue from this product division by the corresponding polynomial. The mathematical notation of this method is as follows:

$$\begin{aligned}
&N_1(x)=b_1(x);\ p_{11}(x)=p_1(x)\bmod p_2(x);\\
&(N_1(x)+\gamma_1(x)p_{11}(x))\bmod p_2(x)=b_2(x);\\
&N_2(x)=N_1(x)+\gamma_1(x)p_1(x);\ p_{12}(x)=p_1(x)p_2(x)\bmod p_3(x);\\
&(N_2(x)+\gamma_2(x)p_{12}(x))\bmod p_3(x)=b_3(x);\\
&N_3(x)=N_2(x)+\gamma_2(x)p_1(x)p_2(x);\ p_{13}(x)=p_1(x)p_2(x)p_3(x)\bmod p_4(x);\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&(N_{i-1}(x)+\gamma_{i-1}(x)p_{1i-1}(x))\bmod p_i(x)=b_i(x);\\
&N_i(x)=N_{i-1}(x)+\gamma_{i-1}(x)p_1(x)p_2(x)p_3(x)\ldots p_{i-1}(x);\ p_{1i}(x)=p_1(x)p_2(x)\ldots p_i(x)\bmod p_{i+1}(x);\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&(N_{k-1}(x)+\gamma_{k-1}(x)p_{1k-1}(x))\bmod p_k(x)=r_k(x);\\
&N(x)=N_k(x)=N_{k-1}(x)+\gamma_{k-1}(x)p_1(x)p_2(x)p_3(x)\ldots p_{k-1}(x).
\end{aligned}\qquad (8)$$

Let us consider an example of polynomial recovery from its residues adding the residue from the product of moduli based on system (7). Since $(x^2+x+1)\bmod(x^2+x+2)\equiv-1$, then from the first comparison (8) we obtain the congruence $(x+3-\gamma_1(x))\bmod(x^2+x+2)\equiv 2x+5$, in which it is necessary to determine a polynomial of the first degree $\gamma_1(x)=A_1x+A_0$ using the method of undetermined coefficients. Combining the last two equalities, it is possible to obtain $A_1=-1, A_0=-2$, respectively $\gamma_1(x)=-x-2$.

At the next stage, it is necessary to find the residue from the product $p_1(x)p_2(x) \bmod p_3(x) = \left((x^2+x+1)(x^2+x+2)\right) \bmod (x^3+2x+1) = (x^4+2x^3+4x^2+3x+2) \bmod (x^3+2x+1) = 2x^2-2x$ and then $N_2(x) = -x^3-3x^2-3x-2$. Thus, $-(x^3+3x^2+3x+2)+\gamma_2(x)(2x^2-2) \bmod (x^3+2x+1) = x^2+4x+1$ or $\gamma_2(x)(2x^2-2) \bmod (x^3+2x+1) = 4x^2+5x+2$. As a result, a polynomial similar to the previous example is obtained.

Therefore, the developed methods for polynomial recovery from its residues in Z[x] ring make it possible to avoid complex operations, in particular, division with residues and finding the inverse element, as well as to perform calculations on polynomials of lower order in comparison with the classical CRT and Garner's algorithm.

Investigating the time complexity of the proposed methods. To calculate the time dependence of the developed methods, it is necessary to determine the most time-consuming basic arithmetic operations. The proposed algorithm includes multiplication, addition, and finding residues in a polynomial ring. In [28], it was proved that the multiplication problem $p(x) \cdot q(x)$ requires $O(4n \log n)$ bit operations (the logarithm is taken at base 2), where $n = \max\{\deg(p(x)), \deg(q(x))\}$ is the largest degree of the polynomial. Taking into account the complexity of finding residues [29], the total estimate is $O(5n \log n)$ of bitwise operations. It should be noted that the developed algorithms require $\frac{(k^2+k)}{2}$ multiplication operations, where k is the number of moduli, as well as finding the residue at each step. Therefore, the total complexity asymptotically approaches $O_1\left(\left(k^2+k\right)n \log n\right)$.

In the classical Garner's algorithm, it is necessary to find the multiplicative inverse k times in the ring of polynomials. In [21] it is stated that the time complexity of the mentioned operation in the standard basis $GF(q^n)$ over $GF(q)$ field, taking into account the complexity of the Euclidean algorithm $O(n \log^2 n)$ and its consequence $O(n \log^2 n \log\log n)$, is equal to $O(n \log^2 n(\log\log n+1))$. In addition, classical Garner's algorithm includes the same operations as the developed algorithms, that is, addition, multiplication, and finding residues. Due to this, its time complexity is $O_2\left((k^2+k)(n \log n)+kn \log^2 n(\log\log n+1)\right)$.

Therefore, the proposed algorithm for number recovery from its residues allows reducing the time complexity from $O_2\left((k^2+k)(n \log n)+kn \log^2 n(\log\log n+1)\right)$ to $O_1\left(\left(k^2+k\right)n \log n\right)$. Figure 1 shows the graphs that indicate the dependences of the time complexities of the proposed and classical approaches to a polynomial recovery from its residues in the ring of polynomials when $k = 10$ and n=1,…,1000.

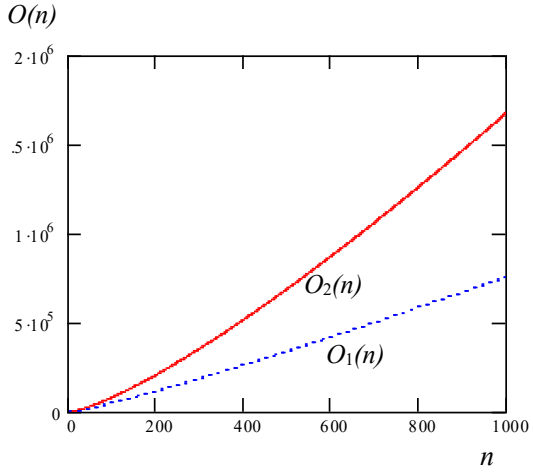


Fig. 1. Graphs of the time complexity dependences of the proposed method $O_1(n)$ and Garner's algorithm $O_2(n)$

It can be seen in Figure 1 that the use of the developed method of a polynomial recovery from its residues in the corresponding ring, which is based on the addition of the product of moduli-polynomials, allows us to reduce the time complexity. The results of the numerical experiment show an increase in both time complexities with an increase in dimensions of the input parameters.

The efficiency of the developed methods is determined by the ratio of the time complexities and in the general case is noted as follows:

$$E(n,k)=\frac{O_2(k,n)}{O_1(k,n)}=\frac{\left(\left(k^2+k\right)\left(n\log gn\right)+kn\log^2 n\left(\log\log gn+1\right)\right)}{\left(\left(k^2+k\right)n\log gn\right)}\approx 1+\frac{\log gn\left(\log\log gn+1\right)}{k+1} \quad (9)$$

Figure 2 shows a graph representing dependence of the proposed method efficiency in comparison with Garner's algorithm. Input parameters are selected within the following ranges: $1\le k\le 20$, a $1\le n\le 1000$.

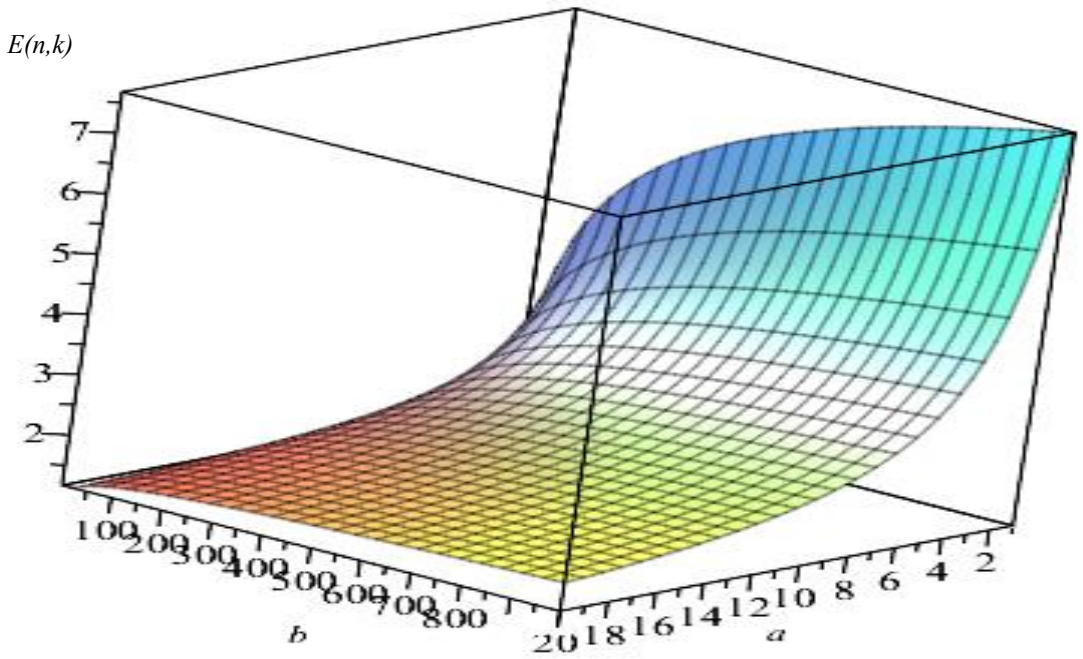


Fig. 2. Dependence of the proposed method efficiency in comparison with Garner's algorithm according to the number of moduli *k* and polynomial degrees *n*

It should be noted that the efficiency of the proposed method increases logarithmically with an increase in the degree of the polynomial, and decreases proportionally with an increase in the number of moduli.

Conclusions. Methods for the polynomial recovery in Z[x] ring are proposed, which make it possible to parallelize calculations and avoid the procedure of finding the polynomial multiplicative inverse due to the operations of adding the product of moduli and the product of residues from moduli that in turn leads to an increase in efficiency. As a result, the effect is achieved when the results of intermediate calculations go beyond the set range, which eliminates the need to find the residue from a polynomial of relatively high order. Analytical expressions of the time complexities of the proposed approach and Garner's algorithm due to the order of polynomials and the number of polynomials-moduli are obtained, which show that the use of the developed method makes it possible to reduce the time complexity. Graphs of the time complexity and efficiency dependences are presented. It is found that the efficiency of the proposed methods increases logarithmically with an increase in the degrees of polynomials, and decreases proportionally with an increase in the number of moduli.

References

1. Milne J.S. Algebraic Number Theory. MilneANT, 2020. 166p. URL: https://www.jmilne.org/math/CourseNotes/ANTc.pdf.
2. Narkiewicz W. Elementary and Analytic Theory of Algebraic Numbers. Berlin. Heidelberg: Springer, 2004. 712 p.
3. Kasianchuk M., Nykolaychuk Ya., Yakymenko I. Theory and Methods of Constructing of Modules System of the Perfect Modified Form of the System of Residual Classes. *Journal of Automation and Information Sciences*. 2016. V.48 (8). P.56-63. DOI: 10.1615/JAutomatInfScien.v48.i8.60.
4. Kasianchuk M., Yakymenko I., Nykolaychuk Y. Symmetric Cryptoalgorithms in the Residue Number System. *Cybernetics and Systems Analysis.* 2021. V. 57(2). P. 329–336. URL: https://doi.org/10.1007/s10559-021-00358-6.
5. Nykolaychuk Ya.M., Yakymenko I.Z.,. Vozna N.Ya, Kasianchuk M.M.. Residue Number System Asymmetric Cryptoalgorithms. *Cybernetics and Systems Analysis*. 2022. V. 58, No. 4, P.611-618. URL: https://doi.org/10.1007/s10559-022-00494-7.
6. Antoniou A., Nakato S., Rissner R. Irreducible polynomials in Int(Z). *ITM Web of Conferences International Conference on Mathematics*. 2018. V. 20, Article Number: 01004. URL: https://doi.org/10.1051/itmconf/20182001004.
7. Ivasiev S., Kasianchuk M., Yakymenko I., Shevchuk R., Karpinski M., Gomotiuk O. Effective algorithms for finding the remainder of multi-digit numbers. *Proceedings of the International Conference on Advanced Computer Information Technology.* 2019. P. 175-178. DOI: 10.1109/ACITT.2019.8779899.
8. Brandon M. M. Polynomial Functions over Rings of Residue Classes of Integers: Thesis. Georgia State University, 2007. 48 p. URL: https://doi.org/10.57709/1059690
9. Bajard J.C., Imbert L., Plantard T. Arithmetic operations in the polynomial modular number system. *Proceedings of the 17th IEEE Symposium on Computer Arithmetic*. 2005. P. 206–213. DOI: 10.1109/ARITH.2005.11.
10. Yatskiv V., Tsavolyk T. Improvement of Data Transmission Reliability in Wireless Sensor Networks on The Basis of Residue Number System Correcting Codes Using the Special Module System. *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON).* 2017. P. 890 893. URL: https://doi.org/10.1109/UKRCON.2017.8100376
11. Su Jun, Yatskiv V. Method and Device for Image Coding & Transferring Based on Residue Number System. *Journal Sensors & Transducers*. 2013. Vol.148. P.60-65.
12. Freed M. RSA Encryption Using Polynomial Rings. Point Loma: Nazarene University< 2018. 17p. URL: https://www.whdl.org/sites/default/files/resource/academic/Freed-RSA%2520Encryption%2520Using%2520Polynomial%2520Rings-HP.pdf
13. Takagi T., Naito S. Construction of RSA Cryptosystem over the Algebraic Field Using Ideal Theory and Investigation of Its Security. *Electronics and Communications in Japan (Part III Fundamental Electronic Science)*. 2000. V. 83 (8). P. 19-29. DOI:10.1002/(SICI)1520-6440(200008)83:83.3.CO;2-S
14. Yakymenko I., Kasianchuk M., Shylinska I., Shevchuk R., Yatskiv V., Karpinski M. Polynomial Rabin Cryptosystem Based on the Operation of Addition. *12th International Conference on Advanced Computer Information Technology (ACIT)*. 2022. P. 345–350. DOI: 10.1109/ACIT54803.2022.9913089.
15. Chu J., Benaissa M. Error detecting AES using polynomial residue number systems. *Microprocessors and Microsystems*. 2013. V. 37. No 2. P. 228-234. URL: https://doi.org/10.1016/j.micpro.2012.05.010
16. Peter R. Turner Residue polynomial systems. *Theoretical Computer Science*. 2002. V.279 (1-2). P. 29–49. URL: https://doi.org/10.1016/S0304-3975(00)00425-4
17. Bajard J.C., Marrez J., Plantard T., Véron P. On Polynomial Modular Number Systems over Z/pZ. *Advances in Mathematics of Communications*. 2022. DOI: 10.3934/amc.2022018. URL: https://arxiv.org/abs/2001.03741

18. Skavantzos A., Stouraitis T. Complex multiplication using the polynomial residue number system: Advances in Computing and Control. *Lecture Notes in Control and Information Sciences*. 1989. V. 130. P. 61-70. URL: https://doi.org/10.1007/BFb0043257h
19. Mahatab K., Sampath K. Chinese remainder theorem for cyclotomic polynomials in Z[X]. *J. Algebra*. 2015. V. 435. P. 223-262. URL: https://doi.org/10.1016/j.jalgebra.2015.04.006.
20. Kalmykov I., Pashintsev V., Tyncherov K., Olenev A., Chistousov N. Error-Correction Coding Using Polynomial Residue Number System. *Appl. Sci.*. 2022. V. *12* (7). P.3365. URL: https://doi.org/10.3390/app12073365
21. Truong T.T., Tran M.T., Duong A.D. Improved Chebyshev Polynomials-Based Authentication Scheme in Client-Server Environment. *Security and Communication Networks*. 2019. V. 2019. Article ID 4250743. 11 p. URL: https://doi.org/10.1155/2019/4250743
22. Lawnik M., Kapczyński A. The application of modified Chebyshev polynomials in asymmetric cryptography. Computer Science. 2019. V. 20 (3). P. 367-381. URL: https://doi.org/10.7494/csci.2019.20.3.3307.
23. Vairachilai S., Kavithadevi M.K., Gnanajeyaraman R. Public Key Cryptosystems Using Chebyshev Polynomials Based on Edge Information. World Congress on Computing and Communication Technologies. 2014. P. 243-245. DOI: 10.1109/WCCCT.2014.21.
24. Li Z.H., Cui Y.D., Xu H.M.. Fast algorithms of public key cryptosystem based on Chebyshev polynomials over finite field. *The Journal of China Universities of Posts and Telecommunications*. 2011. V. 18 (2). P. 86-93. URL: https://doi.org/10.1016/S1005-8885(10)60049-0.
25. Karpinski M., Rajba S., Zawislak S., Warwas K., Kasianchuk M., Ivasiev S., Yakymenko I. A Method for Decimal Number Recovery from its Residues Based on the Addition of the Product Modules. *IEEE 10th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS*). 2019. V.1. P.13–17. DOI: 10.1109/IDAACS.2019.8924395.
26. Nykolaychuk Ya.M., Kasianchuk M.M., Yakymenko I.Z. Theoretical Foundations for the Analytical Computation of Coefficients of Basic Numbers of Krestenson's Transformation. *Cybernetics and Systems Analysis.* 2014. V. 50 (5). P. 649-654. DOI:10.1007/s10559-014-9654-0.
27. Rajba T., Klos-Witkowska AIvasiev., S., Yakymenko I., Kasianchuk M. Research of Time Characteristics of Search Methods of Inverse Element by the Module. IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2017. V.1. P.82–85. DOI: 10.1109/IDAACS.2017.8095054.
28. Harvey D., Hoeven J. Faster polynomial multiplication over finite fields using cyclotomic coefficient rings. *Journal of Complexity*. 2019. V. 54. P.101404. URL: https://doi.org/10.1016/j.jco.2019.03.004.
29. Aho A.V., Hopcroft J.E., Ullman J.D. The Design and Analysis of Computer Algorithms. Wesley Publishing Company, 1974. 480 p.

МЕТОД ВІДНОВЛЕННЯ ПОЛІНОМІВ ЗА ЇХ ЗАЛИШКАМИ НА ОСНОВІ ОПЕРАЦІЇ ДОДАВАННЯ В КІЛЬЦІ Z[X]

І. Якименко, М. Касянчук, І. Шилінська

Західноукраїнський національний університет
11, Львівська вул., Тернопіль, 46009, Україна
Emails: iyakymenko@ukr.net, kasyanchuk@ukr.net

Представлено методи відновлення поліномів за їх залишками в кільці Z[x]. Дана задача є актуальною для застосування в асиметричній та симетричній криптографії, алгоритмах завадозахищеного кодування, контролю та відновленню помилок в процесі передавання пакетів даних в комп'ютерних мережах та розподіленому зберіганню даних. Розглянуто теоретичні основи відновлення поліномів в кільці поліномів на основі відомих підходів, а саме, китайської теореми про залишки, алгоритму Гарнера, визначено їх переваги та недоліки. Розроблено нові методи зворотного перетворення з системи залишкових класів на основі операції додавання добутку модулів та добутку залишків поліномів. Побудовано аналітичні вирази часових складнощів запропонованого методу і алгоритму Гарнера. Представлено їх графічну залежність, яка вказує на те, що розроблений підхід відновлення полінома по його залишках у кільці Z[x] характеризується меншою складністю. Встановлено, що при збільшені розмірності вхідних параметрів зростають часові складності обох методів. Досліджено ефективність використання розробленого методу у кільці Z[x], яка вказує на логарифмічне зростання при збільшенні степенів полінома, і пропорційне зменшення у випадку збільшені кількості модулів.

Ключові слова: відновлення поліномів, кільце поліномів, залишки, Китайська теорема про алгоритм Гарнера, часова складність, ефективність.

DOI 10.15276/imms.v14.no4.314
UDC 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 314-323

INCREASING THE ROBUSTNESS OF THE STEGANOGRAPHIC METHOD WITH CODE CONTROL AND BLIND DECODING

J. K. Ziginova[1], A.V. Sokolov[2]

[1]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
[2]National University "Odesa Law Academy"
23, Fontanska rd, Odesa, 65009, Ukraine
Email: radiosquid@gmail.com

Due to the increasing amount of multimedia content in global traffic, steganographic methods are becoming a key element of information protection systems. Modern steganographic methods fall under such requirements as: ensuring the reliability of perception, resistance to attacks against the embedded message, ensuring sufficient bandwidth. These requirements must be met while ensuring high computational efficiency, which, as practice shows, is possible when performing steganographic transformation in the spatial domain of the container. One of the promising steganographic methods that meet all of the above requirements while ensuring high computational efficiency is the steganographic method with code control, for which a new modification with blind decoding was created. However, this modification was researched only when operating with codewords of size 8x8, while in practice, in the case of creating covert channels in conditions of significant attacks, or when operating with digital video, it may be necessary to ensure greater resistance to attacks against the embedded message. The purpose of this paper is to modify the steganographic method with code control and blind decoding to increase its resistance to attacks against the embedded message by using codewords of size 16x16. The paper shows that simply increasing the length of the codeword does not lead to an increase in the resistance of the steganographic method with code control and blind decoding, which justifies the need to search for new structures of codewords of larger size. The paper proposes a new method of formation of codewords and a modification of the steganographic method with code control and blind decoding based on them. The performed experimental research has allowed us to establish the high efficiency of the proposed modification of the steganographic method, in particular, when using codewords of size 16x16 it becomes possible to ensure an error rate of 1.2% when extracting additional information, which is twice as good as using the original modification of the steganographic method with code control and blind decoding.
Keywords: steganography, code control, Walsh-Hadamard transform, spatial domain.

Introduction and statement of the problem. A significant increase in the amount of multimedia content in global traffic leads to a rise in the relevance of using steganographic methods for information protection. Today, a set of criteria has been formed by which modern steganographic methods are evaluated, among which the following main criteria can be distinguished: ensuring the reliability of perception, high throughput, resistance to attacks against the embedded message, and high computational efficiency. It should be noted that today there are several concepts for building steganographic methods. The first of them involves the use of the transform domain for embedding and extracting additional information, which makes it possible to quite easily provide the specified characteristics of the steganographic message, such as its resistance to attacks against the embedded message. In particular, methods based on the discrete cosine transform [1,2,3,4] have been proposed, which give a high percentage of correctly extracted additional information even under conditions of attacks against the embedded message. For example, a method [1] even under conditions of compression attack with a quality factor QF=70 provides an error rate of 0%. Methods based on singular value decomposition [5,6,7] are also able to provide their high efficiency in countering attacks against

embedded messages. However, the group of methods that use container transform domains for their operation is characterized by a significant drawback, which is associated with the fact that the use of such transforms is characterized by high computational costs both when embedding and when extracting additional information. This significantly limits the use of such methods in many common platforms that are used today, for example, mobile platforms, IoT devices, embedded systems, etc.

In contrast to the methods that use the transform domain for their operation, there is a group of steganographic methods that perform steganographic transformation in the spatial domain of the container. Among such methods, we can distinguish, for example, the classical LSB method and its numerous modifications [8,9,10]. Despite the simplicity of their algorithmic implementation, the absolute majority of steganographic methods operating in the spatial domain of the container are characterized by the inability to resist attacks against the embedded message, in particular compression attacks. In [11], the concept of code control was proposed, which combines the advantages of steganographic methods operating in the transform domains of the container with the computational simplicity offered by steganographic methods operating in the spatial domain of the container. The steganographic method with code control [11] was further developed in [12], where a blind decoding algorithm was proposed, which provides the possibility of extracting additional information that was embedded using the steganographic method with code control without the presence of the original container. In [13], the issue of selecting sets of codewords that provide the best characteristics of the steganographic method with code control and blind decoding was researched. However, to date, only codewords of size 8×8, have been practically researched, while in practice it may be necessary to use codewords of larger size to ensure the greatest robustness of the steganographic method. This is relevant, for example, when creating covert channels, under very strong attacks, or when operating with digital video as a container.

Thus, the relevant task is to research the possibilities and the results provided by the use codewords of size 16×16 in the steganographic method with code control and blind decoding.

The *purpose* of this paper is to modify the steganographic method with code control and blind decoding to increase its resistance to attacks against the embedded message by using codewords of size 16×16.

This paper has the following structure: in Section 2, the original steganographic method with code control and blind decoding is considered. In Section 3, the modified steganographic method for operation with blocks of size 16×16 is proposed. Section 4 presents experimental research and a comparison of the results with other existing methods. Conclusions and suggestions for further work are presented in Section 5.

Modification of the steganographic method with code control and blind decoding. The theoretical basis of the steganographic method with code control [11] is the Walsh-Hadamard transform, which is promising for modern steganographic methods and can be determined using the following relation [14]

$$W_X = H_N^{'} X H_N^{'T}, \tag{1}$$

where $H_N^{'} = \frac{1}{\sqrt{N}} H_N$,

X is matrix of size $N\times N$, the Hadamard matrix H_N is determined by the Sylvester construction

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, \; H_1 = 1. \tag{2}$$

Embedding of additional information in the steganographic method with code control occurs using codewords designed in such a way as to selectively affect a given transformant of the Walsh-Hadamard transform, while according to [11] there is a strict correspondence between the transformants of the Walsh-Hadamard transform and the transformants of the discrete cosine transform. Thus, by selecting the type of codeword, it is possible to embed additional information into a given frequency component of the container, ensuring the given properties of the steganographic message.

Thus, the method is truly resistant to compression attacks, as proven by experiments with embedding information into images. Under conditions of compression attack with a value of QF=70, the error rate value is close to 0 when extracting additional information.

A significant drawback of this method, which limited its practical application, was the lack of the ability to provide blind decoding. This problem was solved in [12] by using spatial and frequency duplication of additional information and the idea of extracting information considering the averaging of subblocks of a macroblock. For the sake of completeness, we will briefly describe this method for the case of using two codewords, which is the most practically valuable.

To embed additional information, the image is divided into blocks of size 8×8, a codeword is added to each block, carrying one bit of additional information

$$M_i = X_i + (-1)^{d_i} * T_8^+, \tag{3}$$

where $T_8^+ = \begin{bmatrix} T_{4_1}^+ + T_{4_2}^+ & T_{4_1}^+ + T_{4_2}^+ \\ T_{4_1}^- + T_{4_2}^- & T_{4_1}^- + T_{4_2}^- \end{bmatrix}$ is the codeword;

X_i is the original image block of size 8×8;

d_i is the bit of additional information;

T_{4_1} and T_{4_2} are the codewords selected for embedding additional information, which exert a concentrated effect on the transformants of the Walsh-Hadamard transform.

The extraction of the additional information is performed using the following steps:

Step 1. The message is divided into blocks M_i' of size $\mu\times\mu$.

Step 2. Each block M_i' of size $\mu\times\mu$ is divided into 4 more blocks of size $\mu/2\times\mu/2$ according to the following construction

$$M_i' = \left[\begin{array}{c|c} \psi_{i11} & \psi_{i12} \\ \hline \psi_{i21} & \psi_{i22} \end{array}\right]. \tag{4}$$

Step 3. For each block M_i', we calculate two matrices $u_{1_{lm}}, u_{2_{lm}}$ of size 2×2 using the following formulas

$$\begin{aligned} u_{1_{lm}} &= \sum_{a=1}^{4}\sum_{b=1}^{4}\psi_{ilm}(a,b)T_{4_1}(a,b); \\ u_{2_{lm}} &= \sum_{a=1}^{4}\sum_{b=1}^{4}\psi_{ilm}(a,b)T_{4_2}(a,b), \quad l,m=1,2. \end{aligned} \tag{5}$$

Step 4. We find the average values

$$\begin{aligned} \overline{u_1} &= \sum_{l=1}^{2}\sum_{m=1}^{2}u_1(l,m); \\ \overline{u_2} &= \sum_{l=1}^{2}\sum_{m=1}^{2}u_2(l,m). \end{aligned} \tag{6}$$

Step 5. We find the value of the extracted additional information bit for this block M_i' as

$$d_i' = \text{sign}((u_{1_{11}} - \overline{u_1}) + (u_{1_{12}} - \overline{u_1}) - (u_{1_{21}} - \overline{u_1}) - (u_{1_{22}} - \overline{u_1}) + \\ + (u_{2_{11}} - u_2) + (u_{2_{12}} - u_2) - (u_{2_{21}} - u_2) - (u_{2_{22}} - u_2)). \quad (7)$$

In [13], all groups of codewords of order 4 were researched. Table 1 shows these codewords and the DCT coefficient mostly affected by the codeword.

Table 1.

Effect of codeword on DCT coefficient

Codeword	DCT transformant	Codeword	DCT transformant
$T_{(1,1)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	(1,1)	$T_{(1,2)} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$	(1,4)
$T_{(1,3)} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$	(1,2)	$T_{(1,4)} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$	(1,3)
$T_{(2,1)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$	(4,1)	$T_{(2,2)} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$	(4,4)
$T_{(2,3)} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$	(4,2)	$T_{(2,4)} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$	(4,3)
$T_{(3,1)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$	(2,1)	$T_{(3,2)} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$	(2,4)
$T_{(3,3)} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$	(2,2)	$T_{(3,4)} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$	(2,3)
$T_{(4,1)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	(3,1)	$T_{(4,2)} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$	(3,4)
$T_{(4,3)} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$	(3,2)	$T_{(4,4)} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$	(3,3)

We present the results of an experiment on increasing the size of the applied codeword to 16×16. Within the framework of this experiment, a similar method of constructing codewords was used as for the size 8×8

$$T_{16}^{+}=\begin{bmatrix} T_{8_1}^{+}+T_{8_2}^{+} & T_{8_1}^{+}+T_{8_2}^{+} \\ T_{8_1}^{-}+T_{8_2}^{-} & T_{8_1}^{-}+T_{8_2}^{-} \end{bmatrix}, \tag{8}$$

where $T_{8_1}^{+}$ and $T_{8_2}^{+}$ are the first and second codeword, respectively.

The use of codewords constructed according to (8) led to the formation of a steganographic message with resistance to attacks against the embedded message comparable to the resistance provided by the method [13], however, with reduced throughput. Such results lead to the conclusion that it is necessary to restructure the codeword with an increase in its size.

The proposed solution. As a basis for constructing codewords T_{16}^{+}, we will use expression (3), accordingly, we will choose the following codeword structure

$$T_{16}^{+}=\begin{bmatrix} T_8^{+} & T_8^{+} \\ T_8^{-} & T_8^{-} \end{bmatrix}=\begin{bmatrix} T_{4_1}^{+}+T_{4_2}^{+} & T_{4_1}^{+}+T_{4_2}^{+} & T_{4_1}^{+}+T_{4_2}^{+} & T_{4_1}^{+}+T_{4_2}^{+} \\ T_{4_1}^{-}+T_{4_2}^{-} & T_{4_1}^{-}+T_{4_2}^{-} & T_{4_1}^{-}+T_{4_2}^{-} & T_{4_1}^{-}+T_{4_2}^{-} \\ T_{4_1}^{-}+T_{4_2}^{-} & T_{4_1}^{-}+T_{4_2}^{-} & T_{4_1}^{-}+T_{4_2}^{-} & T_{4_1}^{-}+T_{4_2}^{-} \\ T_{4_1}^{+}+T_{4_2}^{+} & T_{4_1}^{+}+T_{4_2}^{+} & T_{4_1}^{+}+T_{4_2}^{+} & T_{4_1}^{+}+T_{4_2}^{+} \end{bmatrix}, \tag{9}$$

where $T_{4_1}^{+}$ and $T_{4_2}^{+}$ are the first and second codewords, respectively.

According to the structure of the codeword T_{16}^{+}, we can conclude that the method operates as if one bit of additional information were embedded using four codewords T_8^{+}. Considering the structure of the codeword (9), we write an algorithm for embedding additional information, which is similar to the algorithm for embedding additional information [13].

Step 1. The image is divided into blocks 16×16.

Step 2. A codeword is added to each block, which is modulated by a bit of additional information. Then each subsequent block of the steganographic message M_i will be defined as

$$M_i=X_i+(-1)^{d_i}*T_{16}^{+}, \tag{10}$$

where X_i is the block of the original image of size 16×16;

d_i is the bit of additional information;

T_{16}^{+} is the codeword constructed according to formula (9).

Below we present the algorithm for extracting additional information in the form of specific steps.

Step 1. The original image is divided into blocks of size 16×16 in a standard way. For each block, we define the matrix $P=\begin{bmatrix} p_1 & p_2 \\ p_3 & p_4 \end{bmatrix}$.

Step 2. For each block M_i, we perform its division in a standard way into 4 blocks Y_j, $j=1,2,...,4$ of size 8×8.

Step 3. For each block Y_j of size 8×8, we perform its division into 4 more blocks of size 4×4 according to the following construction $Y_i=\left[\begin{array}{c|c} y_1 & y_2 \\ \hline y_3 & y_4 \end{array}\right]$.

Step 4. For each block Y_j, we calculate the matrices

$$u_{i,1}=\left[\begin{array}{c|c}\sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_1(a,b)T_{4_1}(a,b) & \sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_2(a,b)T_{4_1}(a,b)\\ \hline \sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_3(a,b)T_{4_1}(a,b) & \sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_4(a,b)T_{4_1}(a,b)\end{array}\right],$$
$$u_{i,2}=\left[\begin{array}{c|c}\sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_1(a,b)T_{4_2}(a,b) & \sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_2(a,b)T_{4_2}(a,b)\\ \hline \sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_3(a,b)T_{4_2}(a,b) & \sum\limits_{a=1}^{4}\sum\limits_{b=1}^{4}y_4(a,b)T_{4_2}(a,b)\end{array}\right], \quad (11)$$

where the notation $y(a,b)$ means the element of the matrix with index (a,b).

Step 5. We find the average values

$$\overline{u_{i,1}}=\sum_{l=1}^{2}\sum_{m=1}^{2}u_{i,1}(l,m);\quad \overline{u_{i,2}}=\sum_{l=1}^{2}\sum_{m=1}^{2}u_{i,2}(l,m). \quad (12)$$

Step 6. We find the values $p_i, i=1,2,...,4$ for the given block Y_j as

$$p_i=\text{sign}\left(\sum_{l=1}^{2}\left((u_{i,l}(1,1)-\overline{u_{i,l}})+(u_{i,l}(1,2)-\overline{u_{i,l}})-(u_{i,l}(2,1)-\overline{u_{i,l}})-(u_{i,l}(2,2)-\overline{u_{i,l}})\right)\right). \quad (13)$$

Step 7. We calculate the bit of additional information in the block M_i as $d_j=sign\left(\sum\limits_{l=1}^{4}p_l\right)$.

The results of experiments. For experimental research, 500 images were selected from the NRCS database [16]. Additional information was embedded using the YCbCr space in the Y component of each image block. The pairs $T^{+}_{4(1,4)}$ and $T^{+}_{4(1,4)}$, $T^{+}_{4(2,3)}$ and $T^{+}_{4(3,3)}$, $T^{+}_{4(4,1)}$ and $T^{+}_{4(4,1)}$ were used as codewords T_{4_1} and T_{4_2}.

Experimental research on the stability of the proposed modification of the steganographic method was performed as follows. Additional information was embedded in each image, after which the resulting steganographic message was compressed using the JPEG compression algorithm with a given quality factor *QF*. After that, the embedded information was extracted and the number of errors was estimated.

Fig. 1 shows the results of the experimental research on the number of errors in the extracted additional information depending on the quality factor *QF*.

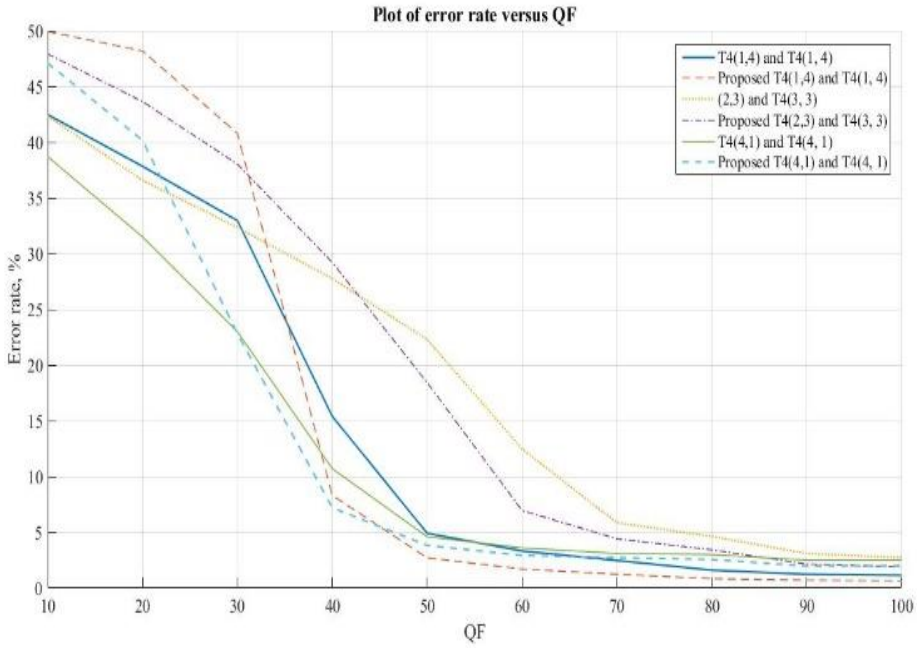


Fig. 1. Number of errors in extracted information depending on the quality factor *QF*

Fig. 2 shows an example of embedding additional information using the proposed modification of the steganographic method.

a) b)

Fig. 2. .Example of embedding additional information using the proposed modification of the steganographic method with code control, a) – original message, b) – steganographic message

Subjective ranking of the images presented in Fig. 2 leads to the conclusion that there are no artifacts or visible distortions in the steganographic message.

Table 2 presents the results of a comparative analysis of the proposed modification of the steganographic method with code control of additional information embedding with the classical steganographic method with code control, the steganographic method with code control and blind decoding, as well as other known analogs.

Table 2.

Comparative analysis of the proposed modification of the steganographic method with code control

Code	QF										PSNR, dB	R	Domain	Blind
	10	20	30	40	50	60	70	80	90	100				
Original [11]														
(5,1)	42.8	29.4	12.2	3.05	0.97	0.72	0.07	0.03	0	0	48.1	1/16	S	-
Modified [13]														
(1,4) (1,4)	42.5	37.8	32.9	15.4	4.9	3.3	2.5	1.6	1.2	1.1	36.1	1/64	S	+
(2,3) (3,3)	42.3	36.6	32.4	27.8	22.3	12.5	5.86	4.63	3.09	2.73	42.2	1/64	S	+
(4,1) (4,1)	38.7	31.5	23.0	10.7	4.6	3.6	3.1	3.0	2.5	2.5	37	1/64	S	+
Proposed														
(1,4) (1,4)	49.9	48.2	40.8	8.3	2.7	1.7	1.24	0.84	0.71	0.65	36	1/256	S	+
(2,3) (3,3)	47.9	43.6	38.0	29.2	18.4	6.96	4.42	3.43	2.15	1.9	36.8	1/256	S	+
(4,1) (4,1)	47.1	40.2	22.8	7.2	3.82	2.93	2.72	2.56	1.98	1.99	36.3	1/256	S	+
[15]														
	-	-	0.02	-	0.02	-	0.01	-	0.01	0.01	27.1	0.01	NN	+
[1]														
	-	-	0	-	0	-	-	0	-	0			DCT	+
[2]														
	-	-	-	-	-	-	33.9	7.4	0.3	-	45	<1/8	DCT	+
[17]														
	13	7	5	4	2	2	2	2	2	-	34.7	1/64	SVD	+
[5]														
	-	-	-	-	23.9	14.1	2.76	0.08	0.08	-	32.7	1/16	SVD	+
[18]														
	-	-	-	-	24.7	14.4	2.71	0.2	0.1	-	32.7	1/64	SVD	+

Analysis of the data presented in Table 2 leads to the conclusion that the proposed modification of the steganographic method with code control allows to provide actually two times fewer errors in the compression attack with the coefficient QF=70 when compared to classical steganographic method with code control and blind decoding. At the same time, on all sets of codewords that were researched, the percentage of correctly extracted additional information is higher than on the same sets of codewords [13]. The PSNR indicator is practically the same as in [13].

Conclusions. The paper proposes a modification of the steganographic method with code control of additional information embedding and blind decoding, which is capable of operating with codewords of size 16×16. The modification, in comparison with analogs, gives better results on the same sets of codewords. At QF values used in real-time information transmission channels, the number of errors is practically 2 times less. The PSNR indicator is constant and equal to ~36 dB, indicating steganographic message reliability of perception.

Further research may concern larger blocks and an increase in the percentage of correctly extracted additional information.

References

1. Wang S., Zheng N., Xu M. A Compression Re-sistant Steganography Based on Differential Manchester Code. *Symmetry*. 2021. V. 13, No. 2. P. 345. URL: https://doi.org/10.3390/sym13020165
2. Zhu Z., Zheng N., Qiao T., Xu M. Robust Ste-ganography by Modifying Sign of DCT Coeffi-cients. *IEEE Access*. 2019. V. 7. P. 168613-168628. URL: https://doi.org/10.1109/access.2019.2953504
3. Rabie T., Kamel I. On the embedding limits of the discrete cosine transform. *Multimed Tools Appl.* 2016. V.75. P.5939–5957. URL:https://doi.org/10.1007/s11042-015-2557-x
4. Rabie T., Kamel I. Toward optimal embedding capacity for transform domain steganography: a quad-tree adaptive-region approach. *Multimed Tools Appl*. 2017. V.76. P.8627–8650. URL:https://doi.org/10.1007/s11042-016-3501-4
5. Chanu Y. J., Singh Kh. M., Tuithung T. A Robust Steganographic Method based on Singular Value Decomposition. *Int. J. of Information & Computation Technology*. 2014. V. 4, No. 7. P. 717-726.
6. Thanki R., Borra S., Dwivedi V., Borisagar K. A steganographic approach for secure communication of medical images based on the DCT-SVD and the compressed sensing (CS) theory. *The Imaging Science Journal*. 2017. V.65., No. 8. P.457-467. URL: https://doi.org/10.1080/13682199.2017.1367129
7. Arunkumar S., Subramaniyaswamy V., Vijayakumar V., Chilamkurti N., Logesh R. SVD-based robust image steganographic scheme using RIWT and DCT forsecure transmission of medical images. *Measurement*. 2019. V.139. P.426-437. URL:https://doi.org/10.1016/j.measurement.2019.02.069
8. Bairagi A. K., Khondoker R., Islam R. An efficient steganographic approach for protecting communication in the Internet of Things (IoT) critical infrastructures. *Information Security Journal: A Global Perspective*. 2016. V.25., No. 4-6. P. 197-212. URL:https://doi.org/10.1080/19393555.2016.1206640
9. Parah S.A., Sheikh J.A., Ahad F., Bhat G.M. High Capacity and Secure Electronic Patient Record (EPR) Embedding in Color Images for IoT Driven Healthcare Systems. *Internet of Things and Big Data Analytics Toward Next-Generation Intelligence*. 2017. V.30. P.411-437. URL:https://doi.org/10.1007/978-3-319-60435-0_17
10. Huang C.T., Tsai M.Y., Lin L.C. VQ-based data hiding in IoT networks using two-level encoding with adaptive pixel replacements. *The Journal of Supercomputing*. 2018. V.74. P.4295–4314. URL: https://doi.org/10.1007/s11227-016-1874-9
11. Kobozeva A.A., Sokolov A.V. Robust Steganographic Method with CodeControlled Information Embedding. *Problemele Energeticii Regionale*. 2021. V.4, No.52. P.115-130. URL: https://doi.org/10.52254/1857-0070.2021.4-52.11
12. Ziginova Yu.K. Modified steganographic method with code control of additional information embedding with blind decoding. *Modern aspects of digitalization and informatization in software and computer engineering*: International scientific and practical conference. 2023. P. 68-70. (In Ukrainian). URL: https://duikt.edu.ua/uploads/n_11337_64054605.pdf
13. Sokolov A.V., Ihnatenko O.O., Balandina N.M. Increasing the Efficiency of Blind Decoding of the Steganographic Method with Code Control of Additional Information Embedding. *Problemele energeticii regionale*. 2024. V.2, No. 62. P.121-137. URL: https://doi.org/10.52254/1857-0070.2024.2-62.11
14. Logachev O. A., Sal'nikov A. A., Jashhenko V. V. Boolean functions in coding theory and cryptology. M.: MCNMO, 2004. 472 p. (In Russian)
15. Li Z., Zhang M., Liu J. Robust image steganography framework based on generative adversarial network. *Journal of Electronic Imaging*. 2021. V. 30, No 2. P. 023006 URL: https://doi.org/10.1117/1.JEI.30.2.023006
16. Natural Resources Conservation Service (NRCS). URL: https://www.nrcs.usda.gov

ПІДВИЩЕННЯ СТІЙКОСТІ СТЕГАНОГРАФІЧНОГО МЕТОДУ З КОДОВИМ УПРАВЛІННЯМ ТА СЛІПИМ ДЕКОДУВАННЯМ

Ю. К. Зігінова[1], А. В. Соколов[2]

[1] Національний університет «Одеська політехніка»
1, Шевченка пр., м.Одеса, 65044, Україна
[2]Національний університет «Одеська юридична академія»
23, Фонтанська дорога, м.Одеса, 65009, Україна
Email: radiosquid@gmail.com

Через зростання долі мультимедійного контенту у світовому трафіку стеганографічні методи стають ключовим елементом систем захисту інформації. До сучасних стеганографічних методів пред'являються такі вимоги як: забезпечення надійності сприйняття, стійкість до атак проти вбудованого повідомлення, забезпечення достатньої пропускної спроможності. Зазначені вимоги мають виконуватися при забезпечення високої обчислювальної ефективності, що, як показує практика, є можливим при виконанні стеганоперетворення у просторовій області контейнера. Одним з перспективних стеганографічних методів, що задовольняє всім зазначеним вимогам при забезпеченні високої обчислювальної ефективності є стеганографічний метод з кодовим управлінням, для якого була створена новітня модифікація із сліпим декодуванням. Тим не менш, зазначена модифікація була досліджена тільки при роботі з кодовими словами розміру 8x8, тоді як на практиці в разі створення прихованих каналів в умовах значних атак, або при роботі з цифровим відео може виникати необхідність забезпечення більшої стійкості до атак проти вбудованого повідомлення. Метою даної статті є модифікація стеганографічного методу з кодовим управлінням та сліпим декодуванням для підвищення його стійкості до атак проти вбудованого повідомлення шляхом застосування кодових слів розміру 16x16. У роботі показано, що просте нарощування довжини кодового слова не призводить до збільшення стійкості стеганографічного методу з кодовим управлінням та сліпим декодуванням, що обґрунтовує необхідність пошуку нових структур кодових слів більшого розміру. У роботі запропоновано новий спосіб формування кодових слів та модифікація стеганографічного методу з кодовим управлінням та сліпим декодуванням на їх основі. Проведені експериментальні дослідження дозволили встановити високу ефективність запропонованої модифікації стеганографічного методу, зокрема, при застосування кодових слів розміру 16x16 стає можливим забезпечити відсоток помилок при вилученні додаткової інформації 1.2%, що є фактично у два рази меншим від застосування оригінальної модифікації стеганографічного методу з кодовим управлінням та сліпим декодуванням.

Ключові слова: стеганографія, кодове управління, перетворення Уолша-Адамара, просторова область.

DOI 10.15276/imms.v14.no4.324
УДК 519.7

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 324-334

МЕТОДИ ПОБУДОВИ ДОСКОНАЛОЇ ФОРМИ СИСТЕМИ ЗАЛИШКОВИХ КЛАСІВ НА МНОЖИНІ ЦІЛИХ КОМПЛЕКСНИХ ЧИСЕЛ

А. М. Алілуйко

Західноукраїнський національний університет
11, Львівська вул., Тернопіль, 46009, Україна
Email: aliluyko82@gmail.com

На даний час велика увага приділяється задачам підвищення швидкодії алгоритмів виконання операцій модулярної арифметики. Досить перспективною для застосування в сучасній теорії чисел, прикладній та обчислювальній математиці, а також асиметричній криптографії, є непозиційна система залишкових класів. Запропонована стаття присвячена розробці методів знаходження набору модулів досконалої форми системи залишкових класів в області цілих комплексних чисел, яке є розширенням множини цілих чисел. Вирішено актуальне завдання знаходження довільної кількості модулів досконалої форми цілочисельної комплексної системи залишкових класів на основі дробових перетворень та факторизації добутку чисел. Використання даного методу дозволяє суттєво зменшити обчислювальну складність при виконанні арифметичних операцій над комплексними числами шляхом розпаралелення процесу обчислень та переведення чисел із системи залишкових класів за рахунок виключення процедури пошуку зворотного елемента за модулем та множенням на базисні числа. Вперше отримано набори трьохмодульної досконалої форми комплексної системи залишкових класів. Визначено умови для знаходження будь-якої кількості модулів, які утворюють досконалу форму комплексної системи залишкових класів, два з яких невідомі. Наведено приклади використання запропонованих методів для досконалої форми системи залишкових класів, у якому отримані всі можливі набори комплексних модулів при заданому найменшому модулі. Представлені табличні значення та проаналізовано графічні залежності норм одержаних модулів. В результаті проведених досліджень показано, що запропонований метод істотно зменшує обчислювальну складність китайської теореми про залишки за рахунок уникнення операції пошуку оберненого елемента за модулем. Використання запропонованого методу підбору модулів, які утворюють досконалу форму, дозволить збільшити швидкодію обчислювальних систем, що працюють у системі залишкових класів.

Ключові слова: система залишкових класів, комплексне число, досконала форма, факторизація, китайська теорема про залишки

Вступ. На сучасному етапі розвитку інформаційних технологій непозиційні системи числення привертають все більшу увагу з метою використання їх при вирішенні ряду проблем та науково-технічних задач [1, 2]. Це пояснюється тим, що при виконанні значних обсягів обчислень в реальному часі суттєво проявляються недоліки двійкової системи (наприклад, наявність міжрозрядних зв'язків та велика розрядність [3]), які суттєво зменшують швидкодію обчислювальних систем [4].

Перераховані недоліки відсутні, наприклад, в системі залишкових класів (СЗК). Зокрема, СЗК ефективно використовується при виконанні цілочисельних операцій модулярної арифметики над багаторозрядними числами: додавання, віднімання, множення, піднесення до степеня [5] і т.д., що особливо актуально в задачах криптографії [4, 6]. Але вона також має певні недоліки (відсутність операцій ділення та порівняння [7], труднощі у виявленні умов переповнення розрядної сітки, складність зворотного перетворення чисел у десяткову систему числення). Безсумнівною перевагою СЗК є можливість виконання операцій над числами, які менші за вибрані модулі, та розпаралелення процесу обчислень. Крім того, використання досконалої (ДФ) [8-10] та модифікованої досконалої форм (МДФ) [11] СЗК дозволяє суттєво

спростити переведення чисел у позиційну системи числення.

Іншим напрямком підвищення швидкодії алгоритмів виконання операцій модулярної арифметики і стійкості комп'ютерних систем до різного виду атак є застосування більш складних структур, зокрема, цілих комплексних чисел або чисел Гауса. Завдяки ізоморфізму між кільцями цілих та комплексних чисел комплексна модулярна арифметика є перспективною для застосування в багатьох системах асиметричної криптографії.

Аналіз досліджень та публікацій. Аналіз наукових досліджень показує, що модулярні операції над цілими комплексними числами можуть успішно застосовуватися в асиметричних криптоалгоритмах RSA, Ель-Гамаля, Рабіна та їх модифікаціях, які раніше були сформульовані для дійсних цілих чисел.

В [12] продемонстровано переваги використання модифікованого алгоритму Ель-Гамаля на множині цілих комплексних чисел щодо підвищення його надійності. В [13, 14] було показано для системи RSA, що можна досягти значного зниження складності обчислень з цілими числами Гауса, але при цьому можливе зниження надійності алгоритму. В [15] зазначено переваги використання гаусівських цілих чисел для генерації відкритого ключа криптосистеми Рабіна. Це дозволило розробити відповідну арифметику для теореми Вільсона і китайської теореми про залишки (КТЗ), а також для обчислення символів Лежандра і квадратичних залишків. Криптографічну еліптичну криву над цілими числами Гауса розглядали в [16]. Крім того, в програмах кодування над цілими числами Гауса використовують цілочисельну арифметику Гауса [17].

При застосуванні комплексної системи залишкових класів (КСЗК) в задачах криптографії доводиться мати справу із подібними труднощами, як і при застосуванні цілочисельної СЗК. Зокрема, при зворотному перетворенні комплексних чисел із КСЗК на основі КТЗ необхідно виконувати громіздку процедуру пошуку оберненого елемента за комплексним модулем [18].

Виходячи із сказаного, актуальною науковою задачею є створення аналітичних методів пошуку комплексних модулів, які задовольняли б умовам ДФ КСЗК.

Мета роботи. Метою даної роботи є подальший розвиток теорії ДФ СЗК. Для досягнення поставленої мети в роботі вирішуються наступні задачі:

– розробка методів побудови ДФ СЗК в комплексній числовій області;

– визначення умов побудови всіх можливих варіантів для заданої кількості модулів ДФ КСЗК.

Основна частина. Теоретичні основи виконання арифметичних модулярних операцій на множині комплексних чисел заклав Гаус [19]. Аналогічно до традиційної асиметричної криптографії, доцільно розглядати тільки цілі комплексні числа (їх ще називають гаусовими), в яких дійсна та уявна частини є цілими. Будь-яке ціле комплексне число $\dot{A} = a + bi,\ a,b \in Z$ записується в СЗК єдиним способом у вигляді набору своїх найменших або абсолютно найменших комплексних залишків $\dot{b}_j$ від ділення $\dot{A}$ на кожен із системи $\dot{M}$ попарно взаємно простих комплексних модулів $\dot{m}_1, \dot{m}_2, \ldots, \dot{m}_n$:

$$\dot{b}_j = \dot{A} \bmod \dot{m}_j, \quad j = \overline{1,n}.$$

При цьому для подання числа $\dot{A}$ в системі $\dot{M}$ найменших залишків, необхідно і достатньо, щоб виконувалися нерівності

$$0 \le ap_M + bq_M < N(\dot{M}),\ \ 0 \le bp_M - aq_M < N(\dot{M}),$$

а в системі $\dot{M}$ абсолютно найменших залишків –

$$\left|ap_M + bq_M\right| \le \frac{1}{2} N(\dot{M}),\ \ \left|bp_M - aq_M\right| \le \frac{1}{2} N(\dot{M}), \qquad (1)$$

де $N(\dot{M})$ – норма комплексного числа $\dot{M} = \prod_{j=1}^{n} \dot{m}_j = p_M + q_M i$.

Відновлення числа $\dot{A}$ з СЗК можна здійснити на основі КТЗ в комплексній числовій області:

$$\dot{A} = \left(\sum_{j=1}^{n} \dot{b}_j \dot{M}_j \dot{f}_j \right) \bmod \dot{M}, \tag{2}$$

де $\dot{M}_j = \dfrac{\dot{M}}{\dot{m}_j}$, $\dot{f}_j$ шукається з виразу $\left(\dot{M}_j \dot{f}_j\right) \bmod \dot{m}_j = 1,\ j = \overline{1,n}$.

Слід зазначити, що пошук обернених мультиплікативних елементів $\dot{f}_j = \dot{M}_j^{-1} \bmod \dot{m}_j$ становить значну обчислювальну складність, який реалізується, наприклад, з допомогою алгоритму Евкліда або використання аналогу функції Ейлера в комплексній числовій області [18]. У роботах [8, 9] було запропоновано ДФ цілочисельної СЗК, у якій здійснювався підбір цілих модулів таким чином, щоб їм відповідали одиничні коефіцієнти $f_j = 1,\ j = \overline{1,n}$.

За аналогією розглянемо побудову ДФ КСЗК, у якій підбір комплексних модулів такий, що

$$\dot{M}_j \bmod \dot{m}_j = 1,\ j = \overline{1,n}, \tag{3}$$

тобто $\dot{f}_j = 1$. Це дозволяє уникнути пошуку обернено елемента і множення в (2) на $\dot{f}_j$. Вираз (2) в цьому випадку спрощується: $\dot{A} = \left(\sum_{j=1}^{n} \dot{b}_j \dot{M}_j \right) \bmod \dot{M}$.

Запишемо вираз (3) у вигляді системи:

$$\begin{cases} \dot{M}_1 \bmod \dot{m}_1 = 1, \\ \cdots \\ \dot{M}_n \bmod \dot{m}_n = 1. \end{cases} \tag{4}$$

Спочатку дослідимо систему з трьох модулів. Для цього задамо модулі $\dot{m}_1$, $\dot{m}_2$ у вигляді: $\dot{m}_2 = a\dot{m}_1 + b$, $\dot{m}_3 = c\dot{m}_1\dot{m}_2 + d = c\dot{m}_1(a\dot{m}_1 + b) + d$, де a і c – натуральні числа, b і d – цілі числа, причому вважаємо, що $N(\dot{m}_1) < N(\dot{m}_2) < N(\dot{m}_3)$. Тоді з (4) отримуємо:

$$\begin{cases} ((a\dot{m}_1 + b)(c\dot{m}_1(a\dot{m}_1 + b) + d)) \bmod \dot{m}_1 = 1, \\ (\dot{m}_1(c\dot{m}_1(a\dot{m}_1 + b) + d)) \bmod (a\dot{m}_1 + b) = 1, \\ (\dot{m}_1(a\dot{m}_1 + b)) \bmod (c\dot{m}_1(a\dot{m}_1 + b) + d) = 1. \end{cases} \tag{5}$$

Третє рівняння системи (5) має місце, якщо, зокрема, добуток $\dot{m}_1(a\dot{m}_1 + b)$ на одиницю більший від значення модуля $c\dot{m}_1(a\dot{m}_1 + b) + d$. Звідси слідує, що $c = 1$, $d = -1$. З другого рівняння бачимо, що $(\dot{m}_1(a\dot{m}_1 + b) - 1) \bmod (a\dot{m}_1 + b) = -1$, тому повинна виконуватися умова $\dot{m}_1 \bmod (a\dot{m}_1 + b) = -1$. Це можливо, якщо $a = 1$, $b = 1$. Тоді система (5) набуде такого вигляду:

$$\begin{cases} ((\dot{m}_1 + 1)(\dot{m}_1(\dot{m}_1 + 1) - 1)) \bmod \dot{m}_1 = 1, \\ (\dot{m}_1(\dot{m}_1(\dot{m}_1 + 1) - 1)) \bmod (\dot{m}_1 + 1) = 1, \\ (\dot{m}_1(\dot{m}_1 + 1)) \bmod (\dot{m}_1(\dot{m}_1 + 1) - 1) = 1. \end{cases} \tag{6}$$

З першого рівняння системи (6) видно, що $(\dot{m}_1(\dot{m}_1 + 1) - 1)) \bmod \dot{m}_1 = -1$. Тоді має виконуватися умова $(\dot{m}_1 + 1) \bmod \dot{m}_1 = -1$ або $1 \bmod \dot{m}_1 = -1$.

Якщо позначити $\dot{m}_1 = x + yi$, то отримаємо таку систему рівнянь

$$\begin{cases} x \bmod (x^2 + y^2) = -x, \\ y \bmod (x^2 + y^2) = -y, \end{cases} \tag{7}$$

яка має розв'язок при $x = \pm 1$, $y = \pm 1$. Отже, вірним є співвідношення $i^k \bmod \dot{m}_1 = i^l$, де $k, l = \overline{0,3}$ при $\dot{m}_1 = \{1+i; 1-i; -1+i; -1-i\}$.

Враховуючи нерівність $N(\dot{m}_1) < N(\dot{m}_2) < N(\dot{m}_3)$, отримуємо такі два єдині набори трьох модулів для ДФ КСЗК: $\{1+i; 2+i; 3i\}$ та $\{1-i; 2-i; -3i\}$.

Метод побудови ДФ КСЗК на основі дробових перетворень. Помноживши кожне рівняння (4) на відповідний модуль, отримаємо:

$$\begin{cases} \dot{M} \bmod \dot{m}_1^2 = \dot{m}_1, \\ \quad \cdots \\ \dot{M} \bmod \dot{m}_n^2 = \dot{m}_n. \end{cases} \tag{8}$$

Використовуючи властивості конгруенцій до системи (8) та КТЗ в комплексній числовій області, матимемо:

$$\dot{M} = \left(\sum_{j=1}^{n} \dot{m}_j \dot{M}_j \dot{f}_j^2 \right) \bmod \dot{P}, \tag{9}$$

де $\dot{P} = \prod_{j=1}^{n} \dot{m}_j^2 = \dot{M}^2$.

Врахувавши, що у ДФ КСЗК $\dot{f}_j = 1$, та скоротивши модуль, ліву та праву частини (8) на їх спільний дільник $\dot{M} = \prod_{j=1}^{n} \dot{m}_j$, перепишемо (9) таким чином:

$$\left(\sum_{j=1}^{n} \dot{M}_j \right) \bmod \dot{M} = 1. \tag{10}$$

Вираз (10) еквівалентний рівності:

$$\sum_{j=1}^{n} \dot{M}_j = \dot{\gamma} \dot{M} + 1, \tag{11}$$

де $\dot{\gamma}$ – ціле комплексне число.

Поділивши ліву та праву частини (11) на $\dot{M}$, отримаємо остаточний вираз для пошуку набору модулів у ДФ КСЗК:

$$\sum_{j=1}^{n} \frac{1}{\dot{m}_j} = \dot{\gamma} + \frac{1}{\prod_{j=1}^{n} \dot{m}_j}. \tag{12}$$

Дослідження рівняння (12) для великої кількості модулів та великого $\dot{\gamma}$ є досить громіздким.

В (12) перенесемо доданок $\frac{1}{m_1}$ вправо. Тоді маємо

$$\frac{1}{\dot{m}_2} + \frac{1}{\dot{m}_3} + \ldots + \frac{1}{\dot{m}_n} = \dot{\gamma} - \frac{1}{\dot{m}_1} + \frac{1}{\dot{m}_1 \dot{m}_2 \dot{m}_3 \ldots \dot{m}_n}. \tag{13}$$

Нехай

$$\dot{\gamma} - \frac{1}{\dot{m}_1} = \frac{s}{\dot{m}_1}, \tag{14}$$

де s – натуральне число, вибір якого дозволяє знайти модуль $\dot{m}_1$ з найменшою нормою.

Знов в (13) перенесемо доданок $\frac{1}{\dot{m}_2}$ вправо. Тоді маємо

$$\frac{1}{\dot{m}_3}+\ldots+\frac{1}{\dot{m}_n}=\dot{\gamma}-\frac{1}{\dot{m}_1}-\frac{1}{\dot{m}_2}+\frac{1}{\dot{m}_1\dot{m}_2\dot{m}_3\ldots\dot{m}_n}. \tag{15}$$

Нехай

$$\dot{\gamma}-\frac{1}{\dot{m}_1}-\frac{1}{\dot{m}_2}=\frac{s}{\dot{m}_1\dot{m}_2}. \tag{16}$$

З (14) та (16) маємо співвідношення:

$$\dot{\gamma}=\frac{s+1}{\dot{m}_1} \text{ та } \dot{m}_2=\frac{s+\dot{m}_1}{s}. \tag{17}$$

Якщо прийняти, що модулі $\dot{m}_1$, $\dot{m}_2$ та $\dot{\gamma}$ є натуральними, то із (17) слідує, що одночасно $\dot{m}_1$ кратне s і $s+1$ кратне $\dot{m}_1$. Це можливо лише при $\dot{m}_1=2$, $s=1$, $\gamma=1$. Такий випадок, детально описано в [10]. Зокрема, отримано ДФ СЗК з цілими модулями при $n=6$ (2, 3, 7, 43, 1807, 3263441, $M=1,0650050423922\times10^{13}$).

Знайдемо можливі комплексні значення $\dot{m}_i$, $\dot{\gamma}$ та натурального s, для яких можуть виконуватися співвідношення (17).

Нехай $\dot{m}_1=x+yi$, $y\neq0$. Тоді $\dot{\gamma}=\frac{s+1}{x+yi}=\frac{s+1}{x^2+y^2}(x-yi)$. Вираз $\frac{s+1}{x^2+y^2}$ повинен мати ціле значення, тому $|s+1|\geq x^2+y^2$.

Друге співвідношення в (17) матиме вигляд $\dot{m}_2=\frac{s+x+yi}{s}=1+\frac{x}{s}+\frac{y}{s}i$. Оскільки $\frac{x^2}{s^2}\geq0$, $\frac{y^2}{s^2}\geq1$, то $x^2+y^2\geq s^2$. Таким чином отримали подвійну нерівність $|s+1|\geq x^2+y^2\geq s^2$, яка виконується лише для одного натурального значення $s=1$ та одного з чотирьох комплексних чисел $\dot{m}_1=\pm1\pm i$ з нормою $N(\dot{m}_1)=2$. Тоді для кожного $\dot{m}_1=\{1+i;1-i;-1+i;-1-i\}$ є відповідне значення $\dot{\gamma}=\frac{2}{\dot{m}_1}=\{1-i;1+i;-1-i;-1+i\}$. Взявши $s=1$, отримаємо $\dot{m}_2=1+\dot{m}_1$.

Якщо в (15) знову перенести доданок $\frac{1}{\dot{m}_3}$ вправо, то з співвідношень $\dot{\gamma}-\frac{1}{\dot{m}_1}-\frac{1}{\dot{m}_2}-\frac{1}{\dot{m}_3}=\frac{1}{\dot{m}_1\dot{m}_2\dot{m}_3}$ та $\dot{\gamma}=\frac{2}{\dot{m}_1}$ $(s=1)$ отримаємо формулу $\dot{m}_3=1+\dot{m}_1\dot{m}_2$. При перенесенні кожного разу в праву частину (11) доданка $\frac{1}{\dot{m}_j}$ вважаємо, що у виразі $\dot{\gamma}-\frac{1}{\dot{m}_1}-\frac{1}{\dot{m}_2}-\frac{1}{\dot{m}_3}-\ldots-\frac{1}{\dot{m}_j}=\frac{s}{\dot{m}_1\dot{m}_2\dot{m}_3\ldots\dot{m}_j}$ завжди $s=1$.

Для останнього модуля $\dot{m}_n$ справедлива рівність

$$\frac{1}{\dot{m}_n}=\frac{1}{\prod_{j=1}^{n-1}\dot{m}_j}+\frac{1}{\dot{m}_n\cdot\prod_{j=1}^{n-1}\dot{m}_j}.$$

Звідси отримуємо, що $\dot{m}_n=\prod_{j=1}^{n-1}\dot{m}_j-1$.

Отже, остаточний вираз для побудови системи модулів ДФ КСЗК при $N(\dot{m}_1) < N(\dot{m}_2) < \ldots < N(\dot{m}_n)$ має такий вигляд:

$$\begin{cases} \dot{m}_1 = 1 \pm i, \\ \dot{m}_k = \prod_{j=1}^{k-1} \dot{m}_j + 1,\ 1 < k < n, \\ \dot{m}_n = \prod_{j=1}^{n-1} \dot{m}_j - 1. \end{cases} \tag{18}$$

Зауважимо, що запропонований в (18) метод не вичерпує всіх можливих наборів модулів ДФ КСЗК при заданих n. Наприклад, при $n = 5$ набір комплексних модулів, отриманий за допомогою (18), буде такий: $1+i$, $2+i$, $2+3i$, $-6+9i$, $-40-117i$. Але можна навести інші набори, зокрема: $1+i$, $2+i$, $2+3i$, $-10+9i$, $4+51i$.

Побудова ДФ КСЗК методом факторизації. Вважаючи, що два останні модулі в рівності (12) невідомі, після відповідних математичних перетворень отримаємо наступне співвідношення:

$$(\dot{m}_{n-1} + \dot{m}_n) \prod_{j=1}^{n-2} \dot{m}_j + \dot{m}_{n-1}\dot{m}_n \prod_{j=1}^{n-2} \dot{m}_j \left(\sum_{k=1}^{n-2} \frac{1}{\dot{m}_k} - \dot{\gamma} \right) = 1. \tag{19}$$

Введемо заміну:

$$\dot{m}_{n-1}, \dot{m}_n = \frac{\dot{a}, \dot{b} - \prod_{j=1}^{n-2} \dot{m}_j}{\prod_{j=1}^{n-2} \dot{m}_j \left(\sum_{k=1}^{n-2} \frac{1}{\dot{m}_k} - \dot{\gamma} \right)}. \tag{20}$$

Після підстановки (20) в (19) отримуємо умову, яка виконується для заданих наборів модулів ДФ КСЗК:

$$\dot{a}\dot{b} = \prod_{j=1}^{n-2} \dot{m}_j \left(\sum_{k=1}^{n-2} \frac{1}{\dot{m}_k} - \dot{\gamma} + \prod_{j=1}^{n-2} \dot{m}_j \right). \tag{21}$$

Для знаходження параметрів $\dot{a}$ і $\dot{b}$ потрібно факторизувати праву частину співвідношення (21). Оскільки модулі $\dot{m}_{n-1}$ та $\dot{m}_n$ цілі комплексні, то з (20) слідують наступні рівності:

$$\left(\dot{a}, \dot{b} - \prod_{j=1}^{n-2} \dot{m}_j \right) \bmod \left(\prod_{j=1}^{n-2} \dot{m}_j \left(\sum_{k=1}^{n-2} \frac{1}{\dot{m}_k} - \dot{\gamma} \right) \right) = 0. \tag{22}$$

Отже, вирази (21) та (22) визначають умови для знаходження довільної кількості модулів ДФ КСЗК, два з яких невідомі.

Нехай $n = 5$, $\dot{\gamma} = 1 - i$ і $\dot{m}_1 = 1+i$, $\dot{m}_2 = 2+i$, $\dot{m}_3 = 2+3i$, отримані з (18). Тоді вирази (20) та (21) будуть мати вигляд:

$$\dot{m}_4 = -7 + 9i - \dot{a},\ \dot{m}_5 = -7 + 9i - \dot{b},\ \dot{a}\dot{b} = -33 - 126i. \tag{23}$$

Права частина останньої рівності в (23) має бути факторизована, на основі чого визначаються параметри $\dot{a}$ та $\dot{b}$. Оскільки $\dot{m}_4$ та $\dot{m}_5$ цілі комплексні числа, то $\dot{a}$ та $\dot{b}$ також мають бути цілими комплексними числами.

Для знаходження $\dot{a}$ та $\dot{b}$ скористаємося властивістю мультиплікативності норми комплексного числа. Тоді має місце розклад:

$$N(\dot{a}) \cdot N(\dot{b}) = N(\dot{a} \cdot \dot{b}) = 16965 = 3 \cdot 3 \cdot 5 \cdot 13 \cdot 29. \tag{24}$$

Використавши всеможливі перестановки множників у (24), можна отримати 12

можливих варіантів наборів для $N(\dot{a})$, $N(\dot{b})$ (табл. 1).

Для знаходження параметра $\dot{a}$ потрібно враховувати, що $N(\dot{a})$ дорівнює сумі квадратів дійсної та уявної частин цілого комплексного числа $\dot{a}$. А з теорії чисел відомо, що числа виду $4k+1$, $k \in N$ можна розкласти в суму двох квадратів. Відповідно до цього можливі значення параметра $\dot{a}$ наведено в таблиці 1.

Таблиця 1.

Можливі значення для $\dot{a}$, $N(\dot{a})$, $N(\dot{b})$

$N(\dot{a})$	$\dot{a}$	$N(\dot{b})$	$N(\dot{a})$	$\dot{a}$	$N(\dot{b})$
1	$\pm 1; \pm i$	$3\cdot3\cdot5\cdot13\cdot29 = 16965$	29	$\pm 2 \pm 5i;$ $\pm 5 \pm 2i$	$3\cdot3\cdot5\cdot13 = 585$
3		$3\cdot5\cdot13\cdot29 = 5655$	$3\cdot13 = 39$		$3\cdot5\cdot29 = 435$
5	$\pm 1 \pm 2i;$ $\pm 2 \pm i$	$3\cdot3\cdot13\cdot29 = 3393$	$3\cdot3\cdot5 = 45$	$\pm 3 \pm 6i;$ $\pm 6 \pm 3i$	$13\cdot29 = 377$
$3\cdot3 = 9$	$\pm 3; \pm 3i$	$5\cdot13\cdot29 = 1885$	$5\cdot13 = 65$	$\pm 1 \pm 8i;$ $\pm 8 \pm i;$ $\pm 4 \pm 7i;$ $\pm 7 \pm 4i$	$3\cdot3\cdot29 = 261$
13	$\pm 2 \pm 3i;$ $\pm 3 \pm 2i$	$3\cdot3\cdot5\cdot29 = 1305$	$3\cdot29 = 87$		$3\cdot5\cdot13 = 195$
$3\cdot5 = 15$		$3\cdot13\cdot29 = 1131$	$3\cdot3\cdot13 = 117$	$\pm 6 \pm 9i;$ $\pm 9 \pm 6i$	$5\cdot29 = 145$

Використовуючи співвідношення в (24) та умову цілочисельності параметра $\dot{b}$, отримуємо 32 можливих наборів з 5 модулів ДФ КСЗК при заданих $1+i$, $2+i$, $2+3i$, які наведені в таблиці 2. Причому, одному і тому ж значенню $N(\dot{a})$ відповідають різні модулі $\dot{m}_4$ та $\dot{m}_5$, що дозволяє змінювати діапазон комплексних чисел.

На рисунку 1 показано характер зміни максимальних та мінімальних значень норм $N(\dot{m}_4)$ та $N(\dot{m}_5)$ залежно від номера норми $N(\dot{a})$, відповідно до таблиці 2 у логарифмічній шкалі. Як видно з рисунка, максимальні значення норм $N(\dot{m}_4)$ зростають, а мінімальні $N(\dot{m}_4)$ спадають приблизно з однаковою інтенсивністю. В той же час, мінімальні значення $N(\dot{m}_5)$ із збільшенням номера норми спадають інтенсивніше, ніж максимальні $N(\dot{m}_5)$.

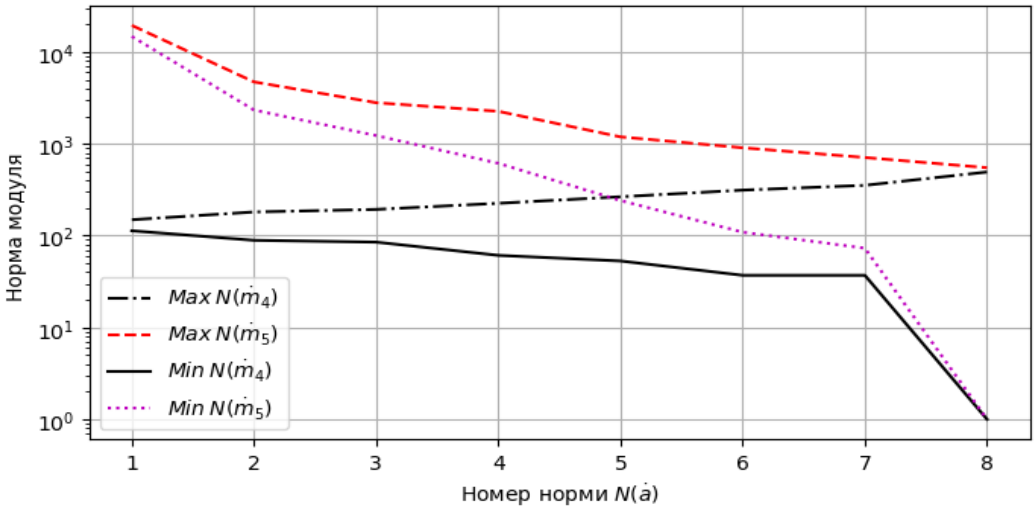


Рис. 1. Характер зміни значень норм $N(\dot{m}_4)$ та $N(\dot{m}_5)$ залежно від номера норми $N(\dot{a})$

Таблиця 2.

Можливі варіанти наборів з 5 модулів ДФ КСЗК при заданих $1+i,\ 2+i,\ 2+3i$

№	$N(\dot{a})$	$\dot{a}$	$\dot{b}$	$\dot{m}_4$	$N(\dot{m}_4)$	$\dot{m}_5$	$N(\dot{m}_5)$
1	1	1	– 33 – 126*i*	– 8 + 9*i*	145	26 + 135*i*	18901
		– 1	33 + 126*i*	–6 + 9*i*	117	– 40 – 117*i*	15289
		i	– 126 + 33*i*	– 7 + 8*i*	113	119 – 24*i*	14737
		– *i*	126 – 33*i*	– 7 + 10*i*	149	– 133 + 42*i*	19453
2	5	1 + 2*i*	– 57 – 12*i*	– 8 + 7*i*	113	50 + 21*i*	2941
		– 1 – 2*i*	57 + 12*i*	– 6 +11*i*	157	– 64 – 3*i*	4105
		2 – *i*	12 – 57*i*	– 9 +10*i*	181	– 19 + 66*i*	4717
		– 2 + *i*	– 12 + 57*i*	– 5 + 8*i*	89	5 – 48*i*	2329
3	9	3	– 11 – 42*i*	– 10 + 9*i*	181	4 + 51*i*	2617
		– 3	11 + 42*i*	– 4 + 9*i*	97	– 18 – 33*i*	1413
		3*i*	– 42 + 11*i*	– 7 + 6*i*	85	35 – 2*i*	1229
		– 3*i*	42 – 11*i*	– 7 + 12*i*	193	– 49 + 20*i*	2801
4	13	2 – 3*i*	24 – 27*i*	– 9 + 12*i*	225	– 31 + 36*i*	2257
		– 2 + 3*i*	– 24 + 27*i*	– 5 + 6*i*	61	17 – 18*i*	613
		3 + 2*i*	– 27 – 24*i*	– 10 + 7*i*	149	20 + 33*i*	1489
		– 3 – 2*i*	27 + 24*i*	– 4 + 11*i*	137	– 34 – 15*i*	1381
5	29	2 + 5*i*	– 24 – 3*i*	– 9 +4*i*	97	17 + 12*i*	433
		– 2 – 5*i*	24 + 3*i*	– 5 + 14*i*	221	– 31 + 6*i*	997
		5 – 2*i*	3 – 24*i*	– 12 + 11*i*	265	– 10 + 33*i*	1189
		– 5 + 2*i*	– 3 + 24*i*	– 2 + 7*i*	53	– 4 – 15*i*	241
6	45	3 + 6*i*	– 19 – 4*i*	– 10 + 3*i*	109	12 + 13*i*	313
		– 3 – 6*i*	19 + 4*i*	– 4 + 15*i*	241	– 26 + 5*i*	701
		6 – 3*i*	4 – 19*i*	– 13 + 12*i*	313	– 11 + 28*i*	905
		– 6 + 3*i*	– 4 + 19*i*	– 1 + 6*i*	37	– 3 – 10*i*	109
7	65	1 – 8*i*	15 – 6*i*	– 8 + 17*i*	353	– 22 + 15*i*	709
		– 1 + 8*i*	– 15 + 6*i*	– 6 + *i*	37	8 + 3*i*	73
		8 + *i*	– 6 – 15*i*	– 15 + 8*i*	289	– 1 + 24*i*	577
		– 8 – *i*	6 + 15*i*	1 + 10*i*	101	– 13 – 6*i*	205
8	117	6 – 9*i*	8 – 9*i*	– 13 + 18*i*	493	– 15 + 18*i*	549
		– 6 + 9*i*	– 8 + 9*i*	– 1	1	1	1
		9 + 6*i*	– 9 – 8*i*	– 16 + 3*i*	265	2 + 17*i*	293
		– 9 – 6*i*	9 + 8*i*	2 + 15*i*	229	– 16 + *i*	257

На рисунку 1 показано характер зміни максимальних та мінімальних значень норм $N(\dot{m}_4)$ та $N(\dot{m}_5)$ залежно від номера норми $N(\dot{a})$, відповідно до таблиці 2 у логарифмічній шкалі. Як видно з рисунка, максимальні значення норм $N(\dot{m}_4)$ зростають, а мінімальні $N(\dot{m}_4)$ спадають приблизно з однаковою інтенсивністю. В той же час, мінімальні значення $N(\dot{m}_5)$ із збільшенням номера норми спадають інтенсивніше, ніж максимальні $N(\dot{m}_5)$.

Застосування ДФ КСЗК у китайській теоремі про залишки. Застосуємо КТЗ до взаємно простих комплексних чисел $\dot{m}_1 = 1+i,\ \dot{m}_2 = 2+i,\ \dot{m}_3 = 2+3i,\ \dot{m}_4 = -6+9i,$ $\dot{m}_5 = -40-117i$, які є основами системи $\dot{M} = \dot{m}_1 \cdot \dot{m}_2 \cdot \dot{m}_3 \cdot \dot{m}_4 \cdot \dot{m}_5 = -12129+9243i$. Спочатку число $\dot{A} = -7+2i$ запишемо в СЗК у вигляді своїх абсолютно найменших комплексних залишків:

$$\dot{b}_1 = \dot{A} \bmod \dot{m}_1 = i,\quad \dot{b}_2 = \dot{A} \bmod \dot{m}_2 = -1,\quad \dot{b}_3 = \dot{A} \bmod \dot{m}_3 = 1+i,$$

$$\dot{b}_4 = \dot{A} \bmod \dot{m}_4 = -1-7i\,,\ \ \dot{b}_5 = \dot{A} \bmod \dot{m}_5 = -7-2i\,.$$

Для $\dot{M}_1 = \dfrac{\dot{M}}{\dot{m}_1} = -1443+10686i$, $\dot{M}_2 = \dfrac{\dot{M}}{\dot{m}_2} = -3003+6123i$, $\dot{M}_3 = \dfrac{\dot{M}}{\dot{m}_3} = 267+4221i$, $\dot{M}_4 = \dfrac{\dot{M}}{\dot{m}_4} = 1333+459i$, $\dot{M}_5 = \dfrac{\dot{M}}{\dot{m}_5} = -39-117i$ обернені елементи за відповідними модулями відомі $\dot{f}_j = \dot{M}_j^{-1} \bmod \dot{m}_j = 1$, $j = \overline{1,5}$. Тут $N(\dot{M}) = 232545690$ і виконуються обмеження (1): $ap_M + bq_M = 103389 < 116272845,\ bp_M - aq_M = 40443 < 116272845$.

Число $\dot{A}$ відновлюється з СЗК за формулою (2):

$$\dot{A} = (i(-1443+10686i) - 1(-3003+6123i) + (1+i)(267+4221i) + (-1-7i)(1333+459i) +$$
$$+(-7-2i)(-39-117i) \bmod(-12129+9243\mathrm{i}) = (-9250-12124i) \bmod(-12129+9243\mathrm{i}) = -7+2i.$$

Отже, число $\dot{A}$ з СЗК відновлюється за допомогою КТЗ без виконання громіздкої операції пошуку оберненого елемента за модулем, а використовуючи операції цілочисельного додавання, множення, та модулярних обчислень в комплексній числовій області.

Висновки. У роботі розв'язано задачу побудови досконалої форми системи залишкових класів на множині цілих комплексних чисел, де відсутня процедура пошуку зворотного елемента по модулю.

Наукова новизна результатів, одержаних у статті, полягає в тому, що вперше запропоновано метод побудови досконалої форми комплексної системи залишкових класів на основі дробових перетворень та факторизації, в якій відсутні операції пошуку оберненого елемента за модулем і множення на базисні числа. Оскільки такі операції характеризуються великою обчислювальною складністю, то отримані методи дозволяють спростити виконання арифметичних операцій над цілими комплексними числами шляхом розпаралелювання процесу обчислень та переведення чисел із системи залишкових класів.

Практична значимість одержаних результатів полягає в тому, що використання запропонованого методу підбору модулів, що утворюють досконалу форму, дозволить збільшити швидкодію обчислювальних систем, що працюють у системі залишкових класів.

Перспективи подальших досліджень полягають у тому, щоб визначити умови для знаходження модулів модифікованої досконалої форми системи залишкових класів, а також програмна та апаратна реалізація запропонованих та запланованих методів.

Список літератури

1. Ananda Mohan P. V. Residue Number Systems: Theory and Applications, Birkhäuser, Basel, 2016. 351 p. URL: http://www.springer.com/978-3-319-41383-9
2. Краснобаєв В., Кошман С., Никольський С., Ковальчук Д. Математична модель надійності комп'ютерної системи у залишкових класах. *Сучасні інформаційні системи*. 2022. Т.6, №4. С. 19–24. URL: https://doi.org/10.20998/2522-9052.2022.4.03
3. Касянчук М., Карпінський М., Казмірчук С. Методологія опрацювання багаторозрядних чисел в асиметричних криптосистемах. *Захист інформації*. 2019. Т.21, №2. С. 65–73. URL: http://dx.doi.org/10.18372/2410-7840.21.13764
4. Adki V., Hatkar S. A Survey on Cryptography Techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2016. V.6, No.6. P. 469–475.
5. Задірака В.К., Терещенко А.М. Комп'ютерна арифметика багаторозрядних чисел у послідовній та паралельній моделях обчислень : монографія. Київ : Наук. думка, 2021. 136 с.

6. Касянчук М.М., Якименко І.З., Івасьєв С.В. Криптосистема Рабіна на основі операції додавання. *Математичне та комп'ютерне моделювання: Технічні науки*. 2019. Вип.19. С.145-150. URL: https://doi.org/10.32626/2308-5916.2019-19.145-150
7. Krasnobayev V.A., Yanko A.S., Koshman S.A. A Method for Arithmetic Comparison of Data Represented in a Residue Number System. *Cybernetics and Systems Analysis*. 2016. V. 52, No. 1. P. 145–150. URL: https://doi.org/10.1007/s10559-016-9809-2
8. Касянчук М. М. Теорія та математичні закономірності досконалої форми системи залишкових класів // Питання оптимізації обчислень. *XXXV Міжнародний симпозіум, Кацивелі*. Київ: Інститут кібернетики ім. В. М. Глушкова. 2009. С. 306–310.
9. Nykolaychuk Ya. M., Kasianchuk M. M., Yakymenko I. Z. Theoretical Foundations for the Analytical Computation of Coefficients of Basic Numbers of Krestenson's Transformation. *Cybernetics and Systems Analysis*. 2014. V. 50, No. 5. P. 649–654. URL: https://doi.org/10.1007/s10559-014-9654-0
10. Касянчук М. М., Якименко І. З., Паздрій І. Р., Николайчук Я.М. Аналітичний пошук модулів досконалої форми системи залишкових класів та їх використання в китайській теоремі про залишки. *Вісник Хмельницького національного університету : технічні науки*. 2015. Т. 221, № 1. С. 170–176. URL: http://journals.khnu.km.ua/vestnik/pdf/tech/2015_1/(221)%202015-1-t.pdf
11. Kasianchuk M., Nykolaychuk Ya., Yakymenko I. Theory and Methods of Constructing of Modules System of the Perfect Modified Form of the System of Residual Classes. *Journal of Automation and Information Sciences*. 2016. V. 48, No 8. P. 56-63. URL: https://doi.org/10.1615/JAUTOMATINFSCIEN.V48.I8.60
12. El-Kassar A., Rizk M., Mirza N., Awad Y. El-Gamal Public-Key Cryptosystem in the Domain of Gaussian Integers. *International Journal of Applied Mathematics*. 2001. V. 7, No. 4. P. 405–412. URL: https://www.researchgate.net/publication/266001078_El-Gamal_public_key_cryptosystem_in_the_domain_of_Gaussian_integers
13. Koval A., Verkhovsky B.S. Analysis of RSA over Gaussian Integers Algorithm. *Fifth International Conference on Information Technology: New Generations (ITNG)*, Las Vegas, NV, USA, 2008. P. 101–105. URL: https://doi.org/10.1109/ITNG.2008.44
14. Koval A. Algorithm for Gaussian Integer Exponentiation. In Information Technology: New Generations. Berlin/Heidelberg: Springer International Publishing, 2016. P. 1075–1085. URL: https://doi.org/10.1007/978-3-319-32467-8_93
15. Awad Y., El-Kassar A.N., Kadri T. Rabin Public-Key Cryptosystem in the Domain of Gaussian Integers. *Proceedings of the International Conference on Computer and Applications (ICCA)*, Beirut, Lebanon. 2018. P. 336–340. URL: https://doi.org/10.1109/COMAPP.2018.8460338
16. Safieh M., Thiers J., Freudenberger J. A Compact Coprocessor for the Elliptic Curve Point Multiplication over Gaussian Integers. *Electronics*. 2020, V. 9, P. 1–21. URL: https://doi.org/10.3390/electronics9122050
17. Rohweder D., Freudenberger J., Shavgulidze S. Low-Density Parity-Check Codes over Finite Gaussian Integer Fields. *Proceedings of the 2018 IEEE International Symposium on Information Theory (ISIT)*, Vail, CO, USA. 2018. P. 481–485. https://doi.org/10.1109/ISIT.2018.8437456
18. Алілуйко А.М., Касянчук М.М. Арифметика асиметричних криптосистем в полі комплексних чисел. *Захист інформації*. 2024. Т. 26, № 1. С. 35–43. URL: https://doi.org/10.18372/2410-7840.26.18825
19. Gauss C. F., Theoria Residuorum Biquadraticorum, Commentatio Secunda, in Werke, Band II. Koniglichen Gesellschaft der Wissenschaften zu Göttingen, 1876. P. 93-148. URL: https://archive.org/details/117771763_002/page/n103/mode/2up

METHODS FOR CONSTRUCTING THE PERFECT FORM OF RESIDUE NUMBER SYSTEM ON THE SET OF COMPLEX INTEGERS

Aliluiko A. M.

West Ukrainian National University
11, Lvivska str., Ternopil, 46009, Ukraine
Email: aliluyko82@gmail.com

So much attention is paid to the tasks of increasing the speed of algorithms for performing modular arithmetic operations. The non-positional residue number system is quite promising for application in modern number theory, applied and computational mathematics, and asymmetric cryptography. This article is focused on the development of methods for finding a set of modules of a perfect-form residue number system in the domain of complex integers, which is an extension of the set of integers. A relevant problem has been solved: finding an arbitrary number of modules of the perfect form of an integer complex residue number system based on fractional transformations and factorization of the product of numbers. The use of this method allows for a significant reduction in computational complexity during arithmetic operations on complex numbers by parallelizing the computation process and converting numbers within the residue number system, eliminating the procedure of finding the inverse element modulo and multiplication by base numbers. Sets of three-module perfect form of the complex residue number system were obtained for the first time. Conditions have been determined for finding any number of modules of modified perfect form of a complex residue number system, with two of them are unknown. Examples of the application of the proposed methods for the perfect form of the residue number system are provided, in which all possible sets of complex modules are obtained for a given smallest module. Tabular values of the obtained modulus norms are presented and their graphical dependencies are analyzed. The results of the conducted research demonstrate that the proposed method significantly reduces the computational complexity of the Chinese Remainder Theorem by avoiding the operation of finding the inverse element modulo. The use of the proposed method for selecting modules that form a perfect form will increase the performance of computational systems operating within the residue number system.

Keywords: residue number system, complex number, perfect form, factorization, the Chinese Remainder Theorem

DOI 10.15276/imms.v14.no4.335
УДК 004.932.1

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 335-343

МЕТОД СЕГМЕНТАЦІЇ МЕТАЛОГРАФІЧНИХ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ НЕЙРОННОЇ МЕРЕЖІ U-NET

Д. Р. Горпенко, Д. М. Кривенко

Національний університет «Одеська політехніка»
1, Шевченко пр., Одеса, 65044, Україна
Emails: horpenko@op.edu.ua, d.krivenko.ilc@gmail.com

У роботі розглянуто завдання сегментації металографічних зображень. Проведено аналіз методів сегментації з використанням нейронних мереж. Запропоновано метод сегментації металографічних зображень з використанням нейронної мережі U-NET, яка демонструє високу якість сегментації навіть за обмеженого розміру навчальної вибірки. Основними процедурами запропонованого методу є попередня обробка, аугментація даних та класифікація об'єктів. Розроблений метод сегментації апробовано на типових металографічних зображеннях. Проведено порівняння результатів сегментації металографічних зображень запропонованим методом та методом вододілів. Оцінку якості сегментації виконано з використанням метрик точності (Accuracy, Precision), повноти (Recall), F-міри та матриці неточностей. На основі експериментальних досліджень визначено, що навчання нейронної мережі U-NET протягом однієї епохи з 500 ітераціями дозволяє отримати найкращі показники якості сегментації, а саме: точність (Accuracy) - 0.91, точність (Precision) - 0.82, повнота (Recall) - 0.87, *F*-міра – 0.85 при достатній оперативності. Метод вододілів забезпечує високу оперативність, однак якість сегментації виявилась нижчою. Таким чином, запропонований в роботі метод є ефективним для задач сегментації металографічних зображень, які потребують високої якості сегментації.
Ключові слова: сегментація; аугментація; нейронна мережа U-NET; метод вододілів, матриця неточностей

Вступ. Контроль якості продукції дозволяє виявити дефекти продукції на ранніх стадіях виробництва та запобігти випуску бракованих виробів з металу. Для проведення контролю якості виробів з металу проводять металографічні дослідження для чого підготовлюють мікрошліфи [1, 2] за допомогою спеціального обладнання [3]. Завдання сегментації металографічних зображень є одним із важливих етапів цифрової обробки зображень в автоматизованих системах контролю якості продукції, що базується на аналізі зображень мікрошліфів цієї продукції [4]. На сьогоднішній день для сегментації зображень не існує загальних методів сегментації, які вирішували б завдання з різних прикладних областей. В [5] автори зазначають, що методи, які застосовують для сегментації металографічних зображень можна розбити на дві групи: методи на основі інтенсивності пікселів або геометричних властивостях форми та методи, засновані на навчанні. До методів першої групи відносяться порогові методи, методи на основі активних контурів. До методів другої групи відносяться методи машинного навчання, які діляться на методи навчання: без вчителя (методи кластеризації, наприклад, метод k -середніх, Principal Component Analysis (PCA)) [6]; з вчителем (методи класифікації, наприклад, нейронні мережі, Support Vector Machine (SVM)) [7]; напівкероване навчання [8]. На сьогоднішній день для сегментації зображень мікрошліфів широко використовують методи із застосуванням нейронних мереж, що забезпечує високу якість сегментації і можуть бути використані в автоматизованих системах контролю якості виробів з металу. Проте використання методів із застосуванням нейронних мереж для сегментації зображень мікрошліфів вимагає значних обчислювальних ресурсів та часу на навчання моделей, що може бути обмеженням для деяких систем, а також вимагає наявності достатнього обсягу

кількості навчальної вибірки. Тому розробка нових методів сегментації, що дозволяють покращити якість сегментації при достатній оперативності та невеликій кількості навчальних даних, є актуальним завданням.

Аналіз літературних джерел та постановка проблеми. Завдання сегментації [9] полягає в розбитті зображення на однорідні за деякою ознакою області. Такими ознаками можуть бути колір, текстура, контури [10]. При використанні методів сегментації зображення в автоматичному режимі можливі такі помилки сегментації зображень: неправильне сегментування - розбіжність контурів сегментованих областей з межами об'єктів на зображенні; пересегментування – збільшення кількості сегментованих областей на зображенні; недосегментування – недостатня кількість сегментованих областей на зображенні. Вибір методу сегментації з урахуванням особливостей зображень – це один із способів зниження помилок при сегментуванні однорідних областей зображення. В результаті застосування методів сегментації одержують безліч однорідних за своєю текстурою областей або безліч контурів. Для різних прикладних завдань розробляються різні методи сегментації, що враховують характеристики об'єктів на зображеннях, що обробляються.

Вище було зазначено, що для сегментації металографічних зображень застосовують методи з допомогою нейронних мереж. Так у роботі [11] використовуються нейронні мережі MLP (Multi Layer Perceptron) та RBF (radial basis function). Складність використання нейронної мережі MLP полягає в тому, що для кожного конкретного завдання необхідно будувати відповідну структуру мережі, що вимагає побудови та налаштування структури мережі. Нейронні мережі RBF забезпечують високу швидкість збіжності. На вхід мережі RBF подається певна кількість вхідних значень та генерується набір вихідних значень, які визначаються вхідними значеннями та набором параметрів, таких як центри мас, вагові значення та інтервалів ширини. Основною складністю при роботі з мережами RBF є вибір цього набору параметрів для навчання мережі.

Останнім часом для вирішення задачі сегментації зображень застосовують згорткову нейронну мережу (CNN - convolutional neural network) [12-14], що забезпечують високу якість сегментації. Однак для якісного навчання мережі потрібна велика кількість зображень.

Тому в роботі пропонується для сегментації металографічних зображень використовувати нейронну мережу U-NET, яка відноситься до CNN. Нейронна мережа U-NET забезпечує високу якість сегментації, при цьому потрібна навчальна вибірка невеликого розміру [15, 16].

Основна частина. Етап сегментації в системах контролю якості продукції дозволяє на металографічних зображеннях виділити зони інтересу, за якими згодом можна виконати оцінку розміру зерна, для дослідження властивостей металу [17]. На рисунку 1 представлено зображення мікроструктури металу.

Рис. 1. Приклад зображення мікроструктури металу

Запропонований метод сегментації металографічних зображень з використанням згорткової нейронної мережі U-NET включає наступні етапи. підготовка навчальної вибірки; налаштування параметрів моделі (розріджена згортка, функція оптимізації

(SGD, Adam); навчання мережі; класифікація об'єктів на зображенні; морфологічна обробка; порогова обробка; оцінка якості сегментації.

Для отримання навченої нейронної мережі була підготовлена навчальна вибірка, що складається з вхідних зображень і масок, що їм відповідають. Маска є картою розмічених класів на навчальних зображеннях [18]. У [19] авторами було описано алгоритм процедури аугментації навчальної вибірки на навчання нейронної мережі U-NET, основними етапами якого є: розтягування і обрізання; поворот на кут 45, 90, 180, 270; зміна яскравості та контрастності; додавання адитивної та імпульсної перешкод.

Для створення масок була запропонована процедура, функціональна схема якої представлена на рисунку 2. На етапі попередньої обробки зменшується рівень адитивного та імпульсного шуму який з'являється на зображеннях мікроструктури сплаву під час отримання зображень. Для зменшення адитивної перешкоди використовувався фільтр Гауса. Для зменшення імпульсного шуму, який виникає під час квантування в процесі отримання цифрового зображення застосовано медіанний фільтр.

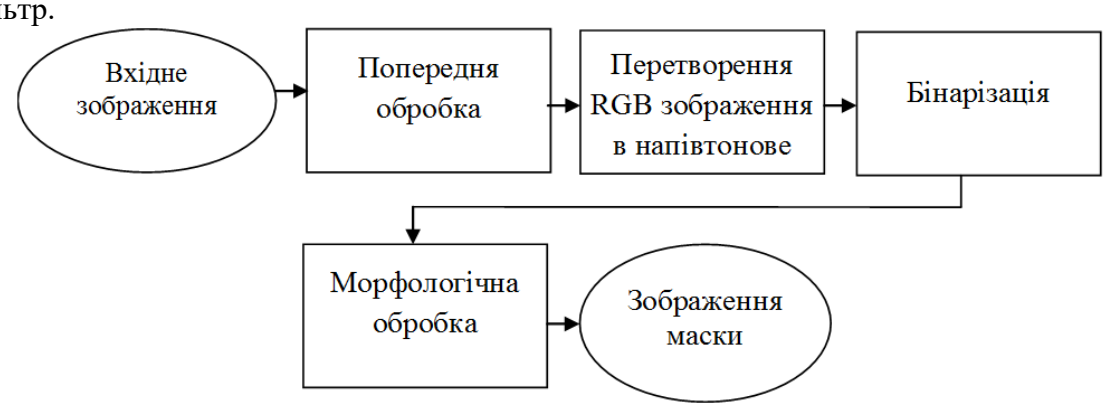


Рис. 2. Функціональна схема процедури створення масок

На етапі бінаризації до напівтонового зображення було застосовано метод Отсу [20] . Морфологічні операції (відкриття, закриття, дилатації, ерозії) [20] застосовувалися для усунення розривів контурів, заповнення областей, усунення дрібних плям на зображеннях після бінаризації.

Основним етапом запропонованого методу є класифікація об'єктів на зображенні за допомогою навченої нейронної мережі U-NET. Мережа U-NET має U - подібну архітектуру [21]. Обробка зображень цією мережею проходить два шляхи: шляху звуження і розширення Звуження (зменшення розмірності зображення) є послідовним виконанням двох операцій згортки 3x3 до зображень з подальшим застосуванням функції активації ReLu (Rectified Linear Unit) [22]. Аналітичний вигляд функція активації ReLu наступний: $A(x) = \max(0, x)$, графік функції активації ReLu показано рисунку 3.

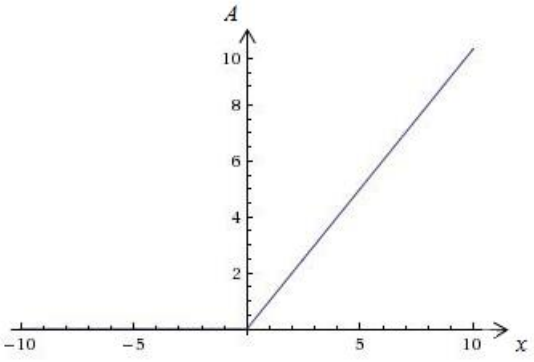


Рис. 3. Графік функції активації ReLu

Застосування функції активації ReLU знижує витрати системних ресурсів та прискорює збіжність стохастичного градієнтного спуску. У результаті на вхід розширюючого шляху надходить інформація про властивості об'єктів, яка збільшується під час проходження через звужуючий шлях.

Обробка на розширюючому шляху полягає у застосуванні згортки 2x2 та об'єднанні з ознаками звужуючого шляху, що містять важливі розрізнювальні характеристики.

Експериментальне дослідження. Програмну реалізацію запропонованого методу сегментації металографічних зображень з використанням нейронної мережі U-NET виконано мовою Python.

Основними програмними модулями є: архітектура мережі (опис шарів); передобробка зображень; навчання мережі; класифікація; постобробка результатів класифікації.

При реалізації програмних модулів використовувалися бібліотеки:

Tkinter та matplotlib - для побудови інтерфейсу програмної реалізації;

Keras та TensorFlow - для побудова нейронної мережі U-NET.

На вхід розробленого програмного модуля надходили зображення мікроструктури металу (табл. 1). Зображення мікроструктури металу були отримані на спеціальному стенді під керівництвом проф. О. Г. Деревянченко.

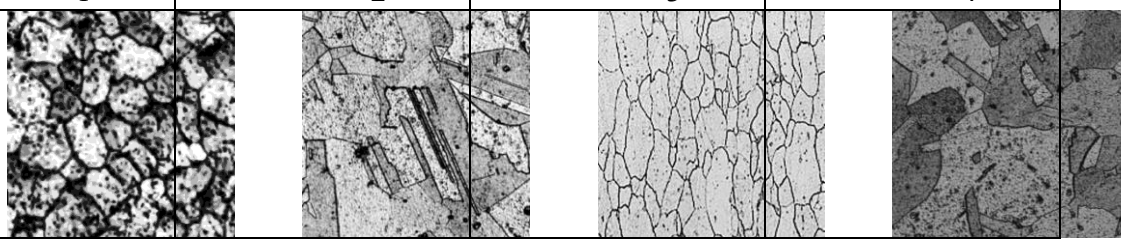
Для зниження рівня адитивного та імпульсного шуму, присутнього на вхідних зображеннях, проводилася попередня обробка.

Фільтр Гауса застосовувався для зменшення адитивного шуму, а медіанний фільтр – для зниження імпульсних перешкод. Після цього виконувалося підсилення контрастності за допомогою еквалізації гістограми.

Вхідні зображення були програмно обрізані до розміру 512x512 пікселів для подачі на вхід мережі U-NET. Приклади тестових зображень наведено в таблиці 1.

Таблиця 1.

Приклади зображень тестової вибірки

1	2	3	4
			

Для порівняння результатів роботи нейронної мережі U-NET було проведено серію експериментів. Спочатку формувалася навчальна вибірка, після чого виконувалося навчання мережі за різної кількості епох та ітерацій.

Також було реалізовано метод вододілів [20] для сегментації зображень мікрошліфів.

Таким чином, було проведено:

Навчання мережі на 1 епосі та 50 ітераціях;

Навчання мережі на 1 епосі та 300 ітераціях;

Навчання мережі на 1 епосі та 500 ітераціях;

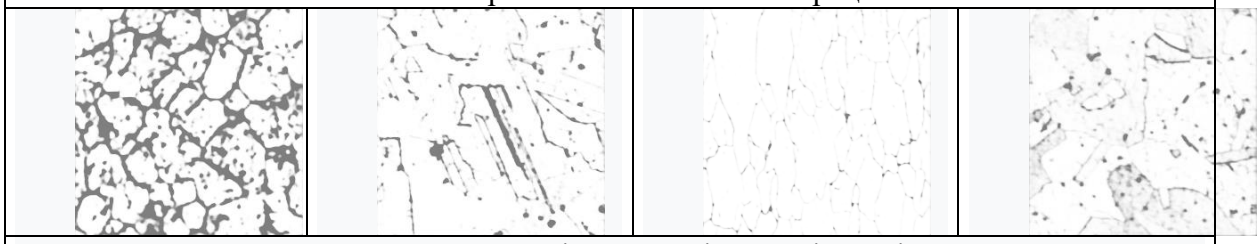
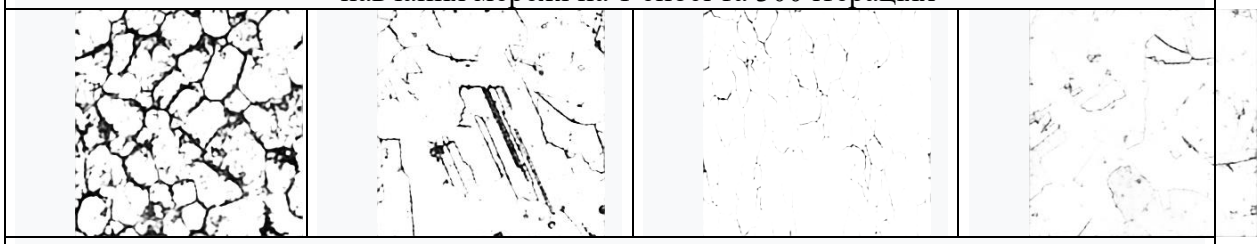
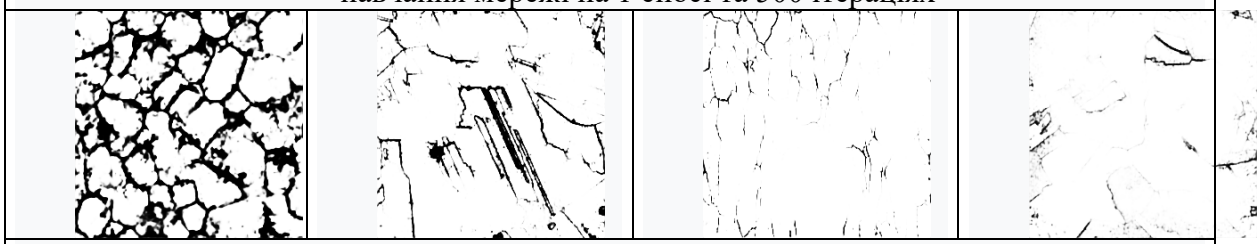
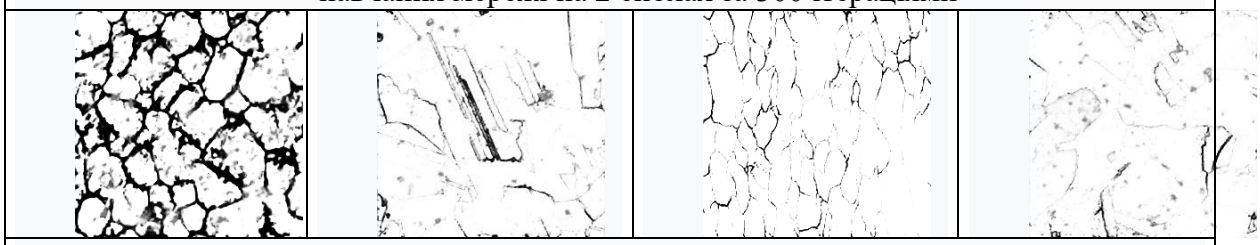
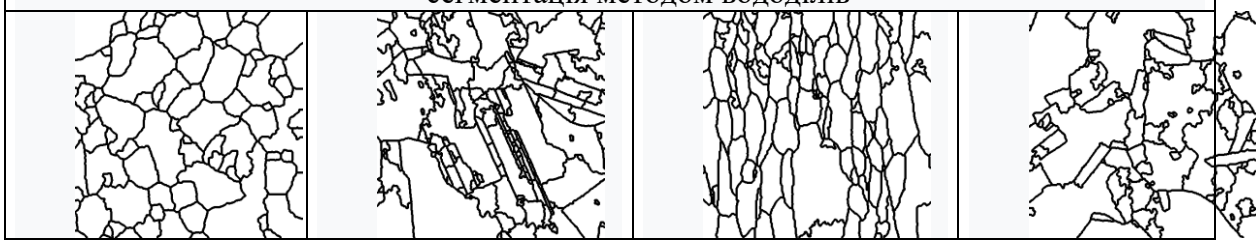
Навчання мережі на 2 епохах за 300 ітераціями;

Сегментація методом вододілів [20] .

Результати роботи мережі після навчання з різною кількістю епох та ітерацій для зображень 1-4 з таблиці 1 представлені в таблиці 2.

Таблиця 2.

Результати тестування мережі

навчання мережі на 1 епосі та 50 ітераціях

навчання мережі на 1 епосі та 300 ітераціях

навчання мережі на 1 епосі та 500 ітераціях

навчання мережі на 2 епохах за 300 ітераціями

сегментація методом вододілів


Як видно з представлення результатів, при навчанні мережі на 1 епосі та 50 ітераціях мережа частково сегментувала зображення, при цьому час навчання було не великим, проте така кількість кроків була занадто малою для отримання хорошого результату сегментації.

При навчанні мережі на 1 епосі та 300 ітераціях якість сегментації значно покращилася завдяки збільшеній кількості кроків, при цьому час навчання збільшився.

При навчанні мережі на 1 епосіта 500 ітераціях якість сегментації покращилася, однак час навчання збільшився практично в два рази.

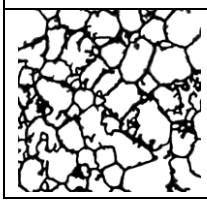
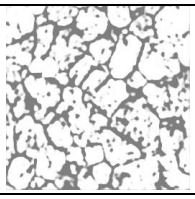
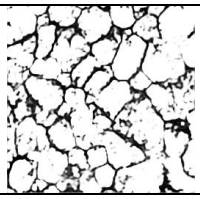
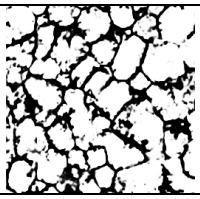
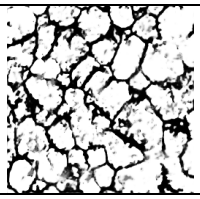
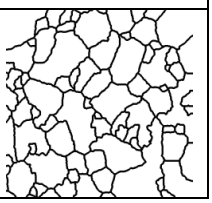
При навчанні мережі на 2 епохах та 300 ітераціях результат виявився близьким до результату при навчанні мережі на 1 епосі та 300 ітераціях.

Для оцінки якості сегментації зображення сегментовані за допомогою методу сегментації з використанням нейронної мережі U-NET при різній кількості епох та

ітерацій, а також методом вододілів, порівнювалися з еталонним зображенням, розміченим експертом (табл. 3).

Таблиця 3.

Порівняння результатів сегментації зображення, розміченого експертом із результатами експерименту

Зображення, розмічене експертом	Розмітка після навчання мережі на 1 епосі та 50 ітераціях	Розмітка після навчання мережі на 1 епосі та 300 ітераціях	Розмітка після навчання мережі на 1 епосі та 500 ітераціях	Розмітка після навчання мережі на 2 епохи за 300 ітераціями	Розмітка після методу вододілів
					

Оцінка якості сегментації зображень проводилася за допомогою матриці неточностей, елементами якої є величини помилок 1-го та 2-го роду (табл. 4) [9].

Таблиця 4.

Загальний вигляд матриці неточностей

Зображення, розмічене експертом	Розмітка зображення, отримана в результаті експерименту	
	Контур 1	Контур 2
Контур 1	TP (true positive)	FP (False positive)
Контур 2	FN (False negative)	TN (true negative)

Елементи матриці неточностей використовувалися для розрахунку оцінок якості сегментації: точності (Accuracy) (1), точності (Precision) (2), повноти (Recall) (3), *F*-міри (4) [9]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

$$P = \frac{TP}{TP + FP}, \quad (2)$$

$$R = \frac{TP}{TP + FN}, \quad (3)$$

$$F_1 = \frac{2 \cdot PR}{P + R}. \quad (4)$$

У задачі сегментації металографічних зображень помилка 1-го роду полягає у прийнятті неправильно класифікованого пікселя за правильний відгук, хоча нульова гіпотеза H_0 стверджує, що піксель на зображенні був класифікований правильно, але ця гіпотеза відкидається. Відповідно, помилка 2-го роду — це неправильно класифіковані пікселі, коли нульову гіпотезу приймають помилково.

У таблиці 5 представлені усереднені оцінки якості сегментації, розраховані для зображень, сегментованих за допомогою розробленого методу сегментації з використанням нейронної мережі U-NET за різної кількості епох та ітерацій, а також методом вододілів.

Таблиця 5.

Оцінки якості сегментації металографічних зображень

Показник якості	після навчання мережі на 1 епосі та 50 ітераціях	після навчання мережі на 1 епосі та 300 ітераціях	після навчання мережі на 1 епосі та 500 ітераціях	після навчання мережі на 2 епохи за 300 ітераціями	після методу вододілів
Accuracy	0,87	0,89	0,91	0,91	0,84
Precision	0,64	0,91	0,82	0,77	0,71
Recall	0,90	0,78	0,87	0,91	0,74
F- міра	0,75	0,84	0,85	0,83	0,72

Проведене експериментальне дослідження показало, що найкращі показники якості сегментації металографічних зображень були отримані при використанні запропонованого методу сегментації з використанням нейронної мережі U-NET, навчання якої проводилося на 1 епосі та 500 ітераціях.

Висновки. У роботі розроблено метод сегментації металографічних зображень із використанням нейронної мережі U-NET. Проведено дослідження якості сегментації зображень під час використання запропонованого методу сегментації з використанням нейронної мережі U-NET з різною кількістю епох та ітерацій, а також методом вододілів. Для збільшення обсягу навчальної вибірки в роботі було виконано процедуру аугментації даних, обсяг якої був збільшений у 10 разів. Крім того, був розроблений набір масок для кожного зображення з навчальної вибірки. Оцінка якості сегментації проводилася з використанням матриці неточності, на основі елементів якої були розраховані показники точності сегментації (Accuracy, Precision, Recall, *F* - міра). Порівняльний аналіз показав, що запропонований метод сегментації з використанням нейронної мережі U-NET, навчання якої проводилося на 1 епосі та 500 ітераціях забезпечило високу якість сегментації. Однак час навчання в цьому випадку зріс практично в 2 рази. Порівняно з методом вододілів, запропонований метод показав більше високу якість сегментації за всіма розглянутими показниками якості сегментації, що є більш пріоритетним, ніж швидкість обробки зображень.

Список літератури

1. Müller M., Stiefel M., Bachmann B.I., Britz D., Mücklich F. Overview: Machine Learning for Segmentation and Classification of Complex Steel Microstructures. *Metals*. 2024. V.14(5). P.553. URL: https://doi.org/10.3390/met14050553
2. Деревянченко А.Г. Некоторые результаты испытаний модуля программного комплекса для обработки изображений микроструктур материалов. *Новые и нетрадиционные технологии в энерго- и ресурсосбережении: материалы международной научно-технической конференции*. Одесса: ОНПУ, 2018. С. 44 – 47.
3. Schubert T., Schneider G., Ketzer-Raichle G., Bernthaler T. The Microstructural Development of Laser-Powder-Bed-Fusion Manufactured Tungsten Carbide–Cobalt Hard Metals. *Practical Metallography*. 2016. V.*53*(7). P.408-421. URL: https://doi.org/10.3139/147.110410
4. Волкова Н.П., Кривенко Д.М. Сегментація металографічних зображень із застосуванням U-NET мережі. *Project, Program, Portfolio Management*: The Proceedings of the International Research Conference, 02 – 03 Desember, 2021, Odesa. Ukraine, P. 123-125.
5. Luengo J., Moreno R., Sevillano I., Charte D., Pelaez-Vegas A., Fernandez-Moreno M., Herrera F. A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and

challenges, *Information Fusion*. 2022. V. 78, P. 232-253. URL: https://doi.org/10.1016/j.inffus.2021.09.018
6. Kunselman C., Sheikh S., Mikkelsen M., Attari V., Arróyave R. Microstructure classification in the unsupervised context. *Acta Materialia*, 2022. V.*223*. P.117434. DOI:10.1016/j.actamat.2021.117434
7. Gola J., Britz D., Staudt T., Winter M., Schneider A. S., Ludovici M, Mücklich F. Advanced microstructure classification by data mining methods. *Computational Materials Science*. 2018. V. 148. P. 324-335. URL: https://doi.org/10.1016/j.commatsci.2018.03.004
8. Chen D., Sun D., Fu J., Liu S. Semi-supervised learning framework for aluminum alloy metallographic image segmentation. *IEEE Access.* 2021. No.9, P.30858-30867. DOI: 10.1109/ACCESS.2021.3059505.
9. Krylov V.N., Volkova N.P. Vector-difference texture segmentation метод in technical and medical express diagnostic systems. *Herald of Advanced Information Technology*. 2020. Vol . 3, No. 4. P. 226 - 239. DOI: 10.15276/hait.04.2020.2
10. Ghahremani M., Ghadiri H., Hamghalam M. Local features integration for content-based image retrieval based on color, texture, and shape. 2021. *Multimed Tools Appl* . V.80. P. 28245–28263. URL: https://doi.org/10.1007/s11042-021-10895-z
11. Papa J.P., de Albuquerque V.H.C., Falcão A.X., Tavares J.M.R.S. (). Fast Automatic Microstructural Segmentation of Ferrous Alloy Samples Using Optimum-Path Forest. *Lecture Notes in Computer Science*. 2010. V. 6026. URL: https://doi.org/10.1007/978-3-642-12712-0_19
12. Wankhade N.P., Sale V.P., Yadav R.S., Jikar P.C., Gadgekar S.R., Dhokey N.B. Metallurgical microstructure classification using CNN: A comprehensive study on heat treatment analysis for steel. *Materials Today: Proceedings*. 2024. URL: https://doi.org/10.1016/j.matpr.2024.05.066.
13. Sun Y., Guan W., Zhang Y. Research on a deep-learning-based method for assessing the metallographic structure of steel used in thermal power plants *Proc. SPIE 13288, Fourth International Conference on Computer Graphics, Image, and Virtualization (ICCGIV)*. 2024. URL: https://doi.org/10.1117/12.3045623
14. Sun Y., Guan W., Zhang Y. Research on a deep-learning-based method for assessing the metallographic structure of steel used in thermal power plants: *Fourth International Conference on Computer Graphics, Image, and Virtualization.* 2024. Vol. 13288. P. 428-435. DOI: 10.1117/12.3045623
15. Motyl M., Madej Ł. Supervised pearlitic–ferritic steel microstructure segmentation by U-Net convolutional neural network. *Archiv.Civ.Mech.Eng.* 2022. V. 22, P. 206. URL: https://doi.org/10.1007/s43452-022-00531-4
16. Polyakova M. V. Image segmentation with a convolutional neural network without pooling layers in dermatological disease diagnostics systems. *Radio Electronics, Computer Science, Control*. 2023. No.1. P. 51-51. DOI: 10.15588/1607-3274-2023-1-5
17. Rusanovsky M., Beeri O., Oren G. An end-to-end computer vision methodology for quantitative metallography. *Sci.Rep.* 2022. V.12. P. 4776. URL: https://doi.org/10.1038/s41598-022-08651-w
18. Пасинков М.К., Хачай М.Ю. Сегментація відбитків пальців з використанням згорткових нейронних мереж. *CEUR Workshop Proceedings*. 2017. V.1894. P.215-225 URL: http://ceurspt.wikidata.dbis.rwth-aachen.de/Vol-1894/mpr3.html
19. Кривенко Д. М. Аугментація навчальної вибірки для навчання U - NET мережі для завдання сегментації металографічних зображень. *Сучасні інформаційні технології*: матеріали X Міжнародної наукової конференції студентів та молодих вчених». Одеса: Наука та техніка, 2020. С. 68-70.
20. Гонсалес Р. Вудс Р. Цифрова обробка зображень. М .: Техносфера , 2005. 1072 с.

21. Siddique N., Paheding S., Elkin C. P., Devabhaktuni V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access*. 2021. V. 9. P. 82031-82057. DOI: 10.1109/ACCESS.2021.3086020.
22. Hara K., Saito D., Shouno H., Analysis of function of rectified linear unit used in deep learning. *International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland*. 2015, P. 1-8. DOI: 10.1109/IJCNN.2015.7280578.

METHOD OF METALLOGRAPHIC IMAGES SEGMENTATION USING THE U-NET NEURAL NETWORK

D. R. Horpenko, D. M. Krivenko

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Emails: horpenko@op.edu.ua, d.krivenko.ilc@gmail.com

This work addresses the problem of metallographic image segmentation. An analysis of segmentation methods using neural networks has been conducted. A segmentation method using the U-NET neural network has been proposed, which demonstrates high segmentation quality even with a limited training dataset size. The main procedures of the proposed method include preprocessing, data augmentation, and object classification. The developed segmentation method has been tested on typical metallographic images. A comparison of segmentation results using the proposed method and the watershed method has been conducted. The segmentation quality has been evaluated using accuracy (Accuracy), precision (Precision), recall (Recall), F-measure, and confusion matrix metrics. Based on experimental studies, it has been determined that training the U-NET neural network for one epoch with 500 iterations yields the best segmentation quality indicators, namely: accuracy (Accuracy) - 0.91, precision (Precision) - 0.82, recall (Recall) - 0.87, F-measure – 0.85, with adequate processing speed. The watershed method ensures high speed, but the segmentation quality is lower. Thus, the proposed method in this work is effective for metallographic image segmentation tasks that require high segmentation quality.
Keywords: segmentation; augmentation; U-NET neural network; watershed method; confusion matrix

DOI 10.15276/imms.v14.no4.344
УДК 004.02:517.958

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 344-349

МЕТОД ГРАНИЧНОЇ КОЛЛОКАЦІЇ ПРИ МОДЕЛЮВАННІ ЗАДАЧ ПРО ВИГИНИ ШАРНІРНО-ОПЕРТОЇ ПЛАСТИНИ З ТОНКИМ ЛІНІЙНИМ ВКЛЮЧЕННЯМ

В. В. Грібова, Л. В. Бовнегра, О. В. Торопенко

Національний університет «Одеська політехніка»
1, Шевченка пр., м. Одеса, 65044, Україна
Emails: gribova@op.edu.ua, dlv5@ukr.net, toropenko.a.v@op.edu.ua

Широке використання конструкцій у вигляді пластин і оболонок в машинобудуванні, кораблебудуванні, приладобудуванні і інших галузях виробництва робить актуальним дослідження міцності і жорсткості елементів цих конструкцій. Наявність в конструкціях підкріплюючих стержнів, опор, прямолінійних дефектів типу тріщин і включень або інших неоднорідностей значно ускладнює їхній розрахунок, так як перераховані елементи в конструкціях є концентраторами напружень. Актуальним є створення нових і удосконалення існуючих методів дослідження концентрації напружень в пружних тілах. Сучасна механіка деформованого твердого тіла послуговується широким набором математичних моделей для опису вигину пластин і оболонок. Із різних моделей контакту пластини з ребром жорсткості найпростішою є модель, в якій передбачається, що ребро жорсткості не має товщини і контакт здійснюється по лінії. Контакт по розімкнутій лінії має місце також у випадку наявності лінійної опори або коли на пластину давить гостролінійний штамп. Всі перераховані випадки (ребро жорсткості, опора, гостролінійний штамп) математично еквівалентні. Математичне формулювання перерахованих випадків призводить до змішаних задач теорії пластин і оболонок, які можуть розглядатись як змішані задачі теорії пружності. Ціллю даної роботи є подальша розробка і деталізація методів рішення задач і дослідження на основі цих методів ряду нових задач пружної рівноваги пластин з лінійним включенням, один або обидва кінця якого знаходяться всередині пластини. Методика, що використовується в роботі полягає в наступному: минаючи стадію зведення задачі до інтегрального рівняння відносно контактних зусиль і рішення цього рівняння, прогини пластини записуються у вигляді лінійної комбінації рішень бігармонійного рівняння в області, яку займає пластина. Невідомі коефіцієнти лінійної комбінації відшукувались методом граничної колокації. Основними результатами роботи є побудова системи рішень бігармонійного рівняння в області з прямолінійним включенням, застосування метода граничної колокації для знаходження коефіцієнтів лінійної комбінації для моделювання систем.

Ключові слова. Система бігармонійних функцій, метод граничної колокації, клас функцій з особливостями, що не інтегруються, вигини пластини, моделювання.

Вступ. Розглядається задача про вигин пластини з тонким лінійним включенням. Відомо, що наявність в конструкціях підкріплюючих стержнів, опор, прямолінійних дефектів типу тріщини і включень значно ускладнює їх розрахунок, так як перераховані елементи в конструкціях є концентраторами напружень. Традиційним методом рішення таких задач є метод, заснований на зведенні задач до інтегральних рівнянь відносно контактних зусиль. Відомі дослідження в цьому напрямку [1 – 4]. Метою роботи є подальша розробка і деталізація методів рішення задач вигину пластин з лінійними включеннями, де рішення задач відбувається у вигляді лінійної комбінації бігармонійних функцій, які враховують наявність включення для кінцевої пластини.

Основна частина. Розглянемо прямокутну шарнірно-оперту пластину ($|x| < a$, $|y| < b$), всередині якої на відрізку $y = 0, |y| \le c$ присутнє тонке жорстке включення. Враховуючи, що поза включенням на пластину не діє розподілене навантаження, приходимо до однорідного бігармонійного рівняння відносно прогинів пластини:

$$\Delta^2\omega(x,y)=0,\quad |x|<a,\ |y|<b \text{ окрім } y=0,|x|<c, \tag{1}$$

де $\omega(x,y)$ – прогини пластини.

Включення переміщується вертикально під дією прикладеного до нього навантаження *P*. В математичній постановці включення можна розглядати як розріз з межами $y=\pm 0,\ |x|\le c$. На межі розрізу виконуються умови:

$$\omega(x,\pm 0)=W_0 \quad |x|\le c, \tag{2}$$

$$\omega_y'(x,\pm 0)=0 \quad |x|\le c, \tag{3}$$

На сторонах пластини задані умови шарнірного спирання:

$$\omega(\pm a,y)=M_x(\pm a,y)=0,\quad |y|\le b, \tag{4}$$

$$\omega(x,\pm b)=M_y(x,\pm b)=0,\quad |x|\le a. \tag{5}$$

Потрібно знайти розподіл прогинів, згинальних моментів, узагальнених перериваючих сил.

З урахуванням парності задачі по *x* і *y* наближене представлення прогину $\omega_N(x,y)$ можна записати у вигляді [5]

$$\omega_N(x,y)=\omega_\Psi(x,y)+\omega_\chi(x,y)+\omega_0(x,y), \tag{6}$$

де $\omega_\Psi(x,y)=\psi_0\,\mathrm{Re}\left\{\left(z\bar{z}+\frac{1}{2}\right)\ln\left(z+\left(z^2-1\right)^{\frac12}\right)-\frac{3}{2}z\left(z^2-1\right)^{\frac12}+2iy\left(z^2-1\right)^{\frac12}\right\}+$

$$+\psi_1\,\mathrm{Re}\left\{\ln\left(z+\left(z^2-1\right)^{\frac12}\right)-z\left(z^2-1\right)^{\frac12}+2iy\left(z^2-1\right)^{\frac12}\right\}+$$

$$+\sum_{n=2}^{N}\psi_n\,\mathrm{Re}\left\{\frac{1}{2(2n-2)}\left(z^2-1\right)^{\frac32}P_{2n-3}^{\frac32,\frac32}(z)-iy\left(z^2-1\right)^{\frac12}P_{2n-2}^{\frac12,\frac12}(z)\right\};$$

$$\omega_\chi(x,y)=\chi_0\,\mathrm{Re}\left\{-(1+\nu)\ln\left(z+\left(z^2-1\right)^{\frac12}\right)-z\left(z^2-1\right)^{\frac12}+(1+\nu)2iy\left(z^2-1\right)^{\frac12}\right\}+$$

$$+\sum_{n=2}^{N}\chi_n\,\mathrm{Re}\left\{\frac{1}{4n}(1+\nu)\left(z^2-1\right)^{\frac32}P_{2n-1}^{\frac32,\frac32}(z)+(1-\nu)iy\left(z^2-1\right)^{\frac12}P_{2n}^{\frac12,\frac12}(z)\right\};$$

$$\omega_0(x,y)=\sum_{n=0}^{N}\left(a_n\,\mathrm{Re}\left(z^{2n}\right)+b_n\,\mathrm{Re}\left(\bar{z}z^{2n+1}\right)\right);$$

$z=x+iy$.

Після заміни багаточленів Якобі $P_n^{\frac12,\frac12}(z)$, $P_n^{\frac32,\frac32}(z)$ багаточленами тих же ступенів від *z,* отримуємо представлення прогину

$$\omega(x,y)=\sum_{n=0}^{N}a_n u_n(x,y); \tag{7}$$

де $u_0(x,y)=\mathrm{Re}\left(\ln\left(z+\left(z^2-1\right)^{\frac12}\right)-z\left(z^2-1\right)^{\frac12}\right);\ u_1(x,y)=\mathrm{Re}\left(\left(z\bar{z}-1\right)\ln\left(z+\left(z^2-1\right)^{\frac12}\right)\right);$

$u_{4n-2}(x,y)=\mathrm{Re}\left(z^{2n-1}\left(z^2-1\right)^{\frac32}\right),\ n=\overline{1,N};\ \ u_{4n-1}(x,y)=\mathrm{Re}\left(2iyz^{2n-2}\left(z^2-1\right)^{\frac12}\right),\ n=\overline{1,N};$

$u_{4n}(x,y)=\mathrm{Re}\,z^{2n-2},\ n=\overline{1,N};\qquad u_{4n+1}(x,y)=\mathrm{Re}\,\bar{z}z^{2n-1},\ n=\overline{1,N}.$

Невідомі постійні коефіцієнти $\psi_n, \chi_n, a_n, b_n \left(n = \overline{0, N}\right)$ шукаються методом граничної колокації [6] – наближеним методом розв'язку диференційних рівнянь, який полягає в зведенні рішення до системи алгебраїчних рівнянь. Для знаходження невідомих постійних коефіцієнтів використовуються граничні умови, які задовольняються не на всьому контурі, а в особливих, попередньо заданих точках колокації. Функція (6) задовольняє рівнянню (1) при будь-яких значеннях 4(N+1) коефіцієнтів $\psi_n, \chi_n, a_n, b_n \left(n = \overline{0, N}\right)$. Легко перевірити, що функція (6) задовольняє умові (3) тотожно. В зв'язку з цим можливі два підходи до рішення задачі (1)… (5):

– шукати наближене представлення прогину у вигляді (6), тоді умова (3) задовольняється тотожно, а умови (2), (4), (5) – колакоційно;

– шукати наближене представлення прогину у вигляді (7), тоді умови (2)… (5) задовольняються в точках колакоційно.

Вибираючи n_1 точок колокації на включенні $y = \pm 0,\ 0 \le x \le c$, n_2 – на стороні $x = a,\ 0 \le y \le b$, n_3 – на стороні $y = b,\ 0 \le x \le a$, і послідовно підставляючи їх значення в (6) і (7), приходимо до системи відповідно $n_1 + 2(n_2 + n_3)$ лінійних алгебраїчних рівнянь в першому випадку і $2(n_1 + n_2 + n_3)$ рівнянь – в другому.

Рішення цих систем, представлене в (6) або (7), дає наближене рішення задачі (1)… (5).

Розрахунки проведені для пластин $a = b$ при відносній довжині включення $\varepsilon = c / a,$ що дорівнює 0,66; 0,5; 0,2; 0,1.Значення прогинів, отримані з допомогою представлень (6) і (7), добре погоджуються між собою.

Прогини для пластини з розмірами $a = b = 2, \varepsilon = 0{,}5$ показані на рис. 1. Прогини максимальні на включенні, де вони рівні $W_0 = 1$, зменшуються до нуля на контурі пластини, що свідчить, що добре задовольняються умовам (2)… (5) обидвома методами.

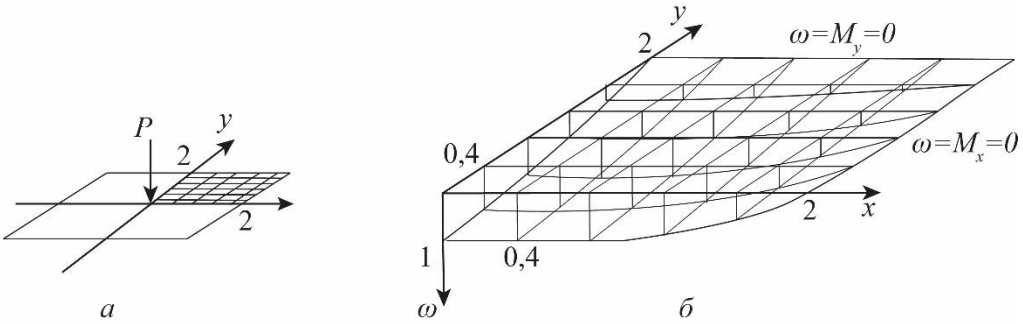


Рис. 1. Розподіл прогинів пластини

Оптимальне з точки зору відповідності граничним умовам, розміщення точок колокації в залежності від величини $\varepsilon = c / a,$ наведено в таблиці 1.

Таблиця 1.

Кількість точок колокації

$\varepsilon = c / a,$	Представлення					
	(6)			(7)		
	n_1	n_2	n_3	n_1	n_2	n_3
0,66	6	5	4	4	3	3
0,5	6	5	4	4	3	3
0,2	3	3	3	4	3	3
0,1	3	3	3	3	3	3

При використанні (6) розраховувалась рівнодіюча контактних зусиль

$$D^{-1}P=\int_0^1\Psi(\xi)d\xi=\psi_0\int_0^1 16\frac{d\xi}{\left(1-\xi^2\right)^{\frac{1}{2}}}+\psi_1\int_0^1 128\frac{P_2^{-\frac{3}{2},-\frac{3}{2}}(\xi)}{\left(1-\xi^2\right)^{\frac{3}{2}}}d\xi+$$

$$+\sum_{n=2}^{N}64n(2n-1)\psi_k\int_0^1\left(1-\xi^2\right)^{-\frac{3}{2}}P_{2n}^{-\frac{3}{2},-\frac{3}{2}}(\xi)d\xi. \qquad (8)$$

Для розрахунку регуляризованих значень розбіжних інтегралів в (8) скористаємось тим, що при $\operatorname{Re}\lambda>-1$ [1, 2].

$$I_n(\lambda)=\int_0^1\left(1-x^2\right)^{\lambda}P_{2n}^{\lambda,\lambda}(x)dx=0\quad(n\geq 0). \qquad (9)$$

Відповідно в регуляризованому значенні $I_n\left(-\frac{3}{2}\right)=0$, звідки $D^{-1}P=8\pi\psi_0$.

При представлені прогину у вигляді (7) рівнодіюча контактних зусиль $D^{-1}P=8\pi a_0$. Ця задача методом інтегральних перетворень зведена до інтегрального рівняння відносно скачка рівнодіючої контактних зусиль [1, 2]. Порівняння отриманих значень безрозмірного коефіцієнта $\alpha=10^3\left(Pa^2\right)^{-1}DW$ (6), (7) з відомими [2, 3] наведено в табл. 2.

Таблиця 2.

Значення безрозмірного коефіцієнта α

$\varepsilon=c/a,$	$\alpha=10^3\left(Pa^2\right)^{-1}DW$ (6), (7)		
	[1, 2]	(6)	(7)
0,66	2,16	2,29	2,13
0,5	4,40	4,64	4,68
0,2	9,60	9,63	9,17
0,1	10,75	10,98	10,60

Графіки згинальних моментів M_x, M_y вздовж лінії $y=\pm 0,\ 0\leq x\leq a$ наведені на рис. 2; a = 2. При $(x,y)\to(1,0)$ $M_x, M_y\to\infty$ як $r^{\frac{1}{2}}$, $r=\sqrt{(x-1)^2+y^2}$.

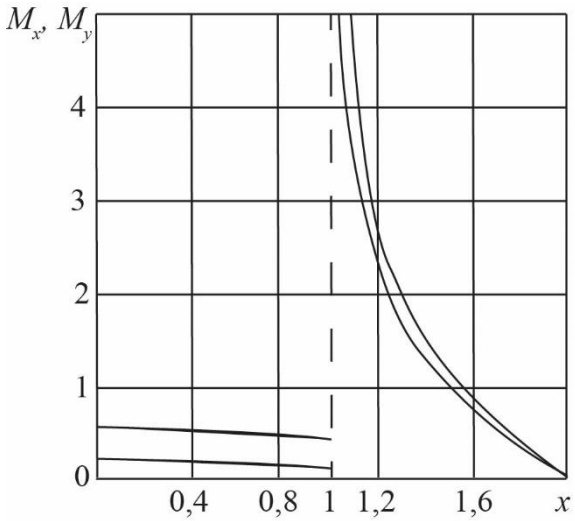


Рис. 2. Графіки згинальних моментів M_x, M_y вздовж лінії $y=\pm 0,\ 0\leq x\leq a: D=1, P=13{,}9$.

Епюри величин $K_x = \lim \frac{M_x}{P}\frac{\sqrt{r}}{\sqrt{c}}$, $K_y = \lim \frac{M_y}{P}\frac{\sqrt{r}}{\sqrt{c}}$, при $r = 10^{-4}$ наведені на рис.3. Якісна картина аналогічна результатам [3], де розглядались безкінечні пластини.

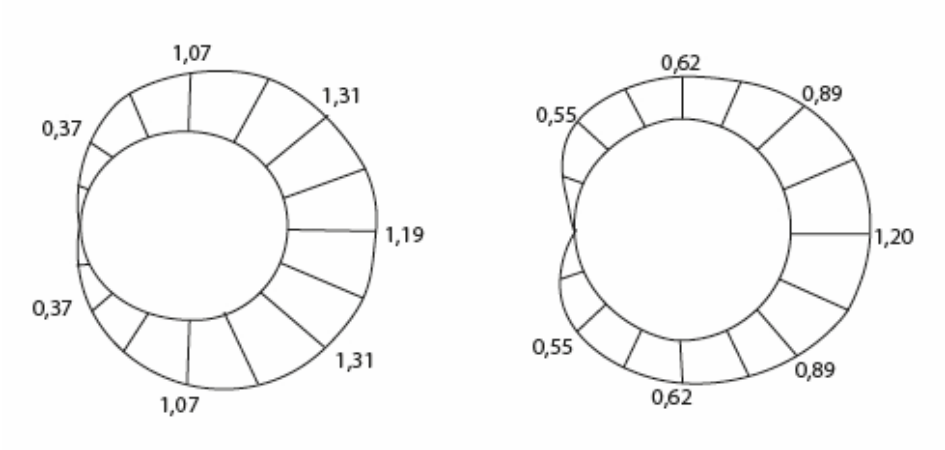


Рис. 3. Епюри $K_x(a)$, $K_y(b)$.

Отримані результати показують, що використання методу граничної колокації достатньо ефективне при вирішенні задач вигину нескінчених пластин з включеннями.

Список літератури

1. Кривий О. Ф., Морозов Ю. О. Особливості поля напружень в околі міжфазного кругового включення при змішаних умовах контакту із кусково-однорідним трансверсально-ізотропним простором. *Прикладна механіка*. 2024. Т. 60, № 3. С. 91–100. URL: http://jnas.nbuv.gov.ua/article/UJRN-0001493980
2. Usov A., Morozov Yu., Kunitsyn M., Tonkonozhenko A., Chernush I.. Investigation of the influence of structural inhomogeneities on the strength of welded joints of functionally gradient materials. *Proceedings of Odessa Polytechnic University.* 2020. No. 1. V. 60.
3. DOI: https://doi.org/10.15276/opu.1.60.2020.03
4. Kryvyi O.F., Morozov Y. O. Influence of Concentrated Forces on an Interface Inclusion under the Conditions of Smooth Contact in the Inhomogeneous Transversely Isotropic Space. *J Math Sci* 279. 2024. P.197–212. DOI: https://doi.org/10.1007/s10958-024-07005-3
5. Kryvyi O. F., Morozov Y. O. Stress Field Features in Vicinity of Interfacial Circular Inclusion Under Mixed Contact Conditions with Piecewise Homogeneous Transversely Isotropic Space. *Int Appl Mech.* 2024. V.60, P.331–340. DOI: https://doi.org/10.1007/s10778-024-01286-6
6. Грібова В. В., Пєрстньова В. В. Розрахунок деформації прямокутної пластини з тонким лінійним включенням. *Actual problems of practice and science methods of their solution*: *The IV International Science Conference*. 2022. P. 515–518.

THE METHOD OF BOUNDARY COLLOCATION FOR SIMULATING PROBLEMS ABOUT CURVES OF A HINGED-OPERATED PLATE WITH A THIN LINEAR INCLUSION

V. Gribova, L. Bovnegra, O. Toropenko

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Emails: gribova@op.edu.ua, dlv5@ukr.net, toropenko.a.v@op.edu.ua

The widespread use of structures in the form of plates and shells in mechanical engineering, shipbuilding, instrumentation and other industries makes it important to study the strength and stiffness of the elements of these structures. The presence of reinforcing rods, supports, straight-line defects such as cracks and inclusions or other inhomogeneities in structures significantly complicates their calculation, as these elements in structures are stress concentrators. It is important to develop new and improve existing methods for studying stress concentration in elastic bodies. Modern mechanics of deformable solids uses a wide range of mathematical models to describe the bending of plates and shells. Of the various models of plate contact with a stiffener, the simplest is the model in which the stiffener is assumed to have no thickness and contact is made along a line. Open line contact also occurs when there is a linear support or when the plate is pressed against a sharp-edged die. All of these cases (stiffener, support, sharp-line stamp) are mathematically equivalent. The mathematical formulation of these cases leads to mixed problems of the theory of plates and shells, which can be considered as mixed problems of the theory of elasticity. The purpose of this paper is to further develop and detail the methods for solving the problems and to study, on the basis of these methods, a number of new problems of elastic equilibrium of plates with a linear inclusion, one or both ends of which are inside the plate. The methodology used in this work is as follows: bypassing the stage of reducing the problem to an integral equation with respect to contact forces and solving this equation, the deflections of the plate are written as a linear combination of solutions of the biharmonic equation in the region occupied by the plate. The unknown coefficients of the linear combination were found by the boundary collocation method. The main results of the work are the construction of a system of solutions for a biharmonic equation in the domain with a straightforward inclusion, and the application of the method of boundary collocation to find the coefficients of the linear combination for simulating systems.

Keywords: system of biharmonic functions, boundary collocation method, class of functions with noninte-grated singularities, deflections of a plate, simulation.

DOI 10.15276/imms.v14.no4.350
УДК 004.42:004.272:519.612.2

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 350-357

КРОСПЛАТФОРМЕНА СИСТЕМА АНАЛІЗУ ЕФЕКТИВНОСТІ ПАРАЛЕЛЬНИХ ОБЧИСЛЮВАЛЬНИХ АЛГОРИТМІВ

О. О. Жульковський[1], Г. Я. Вохмянін[1], І. І. Жульковська[2],
Ю. В. Ульяновська[2], В. А. Рябоволенко[2]

[1]Дніпровський державний технічний університет
2, Дніпробудівська вул., м. Кам'янське, 51918, Україна
[2]Університет митної справи та фінансів
2/4, Володимира Вернадського вул., м. Дніпро, 49000, Україна
Email: olalzh@ukr.net

Представлено результати розроблення та дослідження спеціалізованої системи для автоматизованого порівняльного аналізу ефективності обчислювальних алгоритмів, зокрема з використанням технології SIMD. Основною метою дослідження є створення кросплатформеного програмного забезпечення для проведення обчислювальних експериментів, що дозволяє порівнювати продуктивність стандартних та оптимізованих реалізацій алгоритмів розв'язання систем лінійних алгебраїчних рівнянь. Дослідження зосереджено на підходах до розроблення та порівнянні продуктивності реалізації фундаментальних методів розв'язання, зокрема методів Гауса та спряжених градієнтів. Для досліджуваних методів розроблено по дві версії реалізації – стандартна та оптимізована з використанням SIMD-інструкцій. Програмна реалізація алгоритмів виконана продуктивними засобами Microsoft Visual Studio C++ із використанням стандартної спеціалізованої бібліотеки immintrin.h та набору команд процесора AVX. Розроблений програмний комплекс базується на сучасному стеку технологій, зокрема використовує фреймворк Nextron, який поєднує можливості Electron для створення кросплатформених застосунків та Next.js для побудови інтерактивного користувацького інтерфейсу. Архітектура системи забезпечує модульність, масштабованість та зручність використання завдяки застосуванню компонентного підходу React. У роботі представлено результати експериментального дослідження ефективності використання SIMD-технології для пришвидшення обчислювальних алгоритмів. Проведені експерименти демонструють суттєве підвищення продуктивності реалізації паралельних обчислювальних алгоритмів у діапазоні 2.67–5.81 разів у залежності від розмірності СЛАР та обраного обчислювального методу. Виявлено, що ефективність SIMD-оптимізації зростає пропорційно збільшенню обсягу вхідних даних, тоді як для малих розмірностей накладні витрати на векторизацію можуть лімітувати загальну продуктивність. Результати дослідження підтверджують потенціал у використанні SIMD-технології для оптимізації обчислювальних алгоритмів та демонструють практичну цінність розробленого програмного забезпечення для автоматизації процесу проведення та аналізу обчислювальних експериментів. Отримані результати можуть бути використані при розробці інших оптимізованих алгоритмів та програмних систем, зокрема для комп'ютерного моделювання, орієнтованих на високопродуктивні обчислення.

Ключові слова: кросплатформений застосунок, SIMD, паралелізм на рівні даних, пришвидшення обчислень, immintrin.h, Electron, Next.js, Nextron.

Вступ. Наукові дослідження, пов'язані із проведенням обчислювальних експериментів, потребують значних часових витрат на підготовку та аналіз експериментальних даних, а процес систематизації та візуалізації отриманих результатів часто виконується вручну з використанням непов'язаних між собою інструментів. Такі підходи мають низьку ефективність із-за збільшення часу проведення досліджень, високої ймовірності виникнення помилок під час обробки даних, залучення додаткових ресурсів тощо. У цьому контексті розроблення спеціалізованого застосунку для автоматизації процесу

проведення обчислювальних експериментів є актуальним завданням, вирішення якого дозволить створити проблемно-орієнтоване програмне забезпечення не тільки для проведення експериментів, а й автоматизованої обробки результатів та їх візуалізації у вигляді інформативних графіків та таблиць.

SIMD (Single Instruction, Multiple Data – одиночний потік команд, множинний потік даних) є одним із сучасних інструментів для пришвидшення виконання алгоритмів, які працюють з великими обсягами однорідних даних, шляхом одночасної обробки декількох елементів у межах однієї інструкції [1]. Порівняння продуктивності класичних алгоритмів і їх оптимізованих версій дозволяє визначити не лише швидкість виконання, але й вплив різних факторів, таких як розмір вхідних даних, конфігурація обладнання та оптимізація коду [2].

Окремі бібліотеки мови програмування JavaScript формують сучасний стек технологій для розроблення кросплатформеного застосунку з підтримкою вебтехнологій [3]. Їх поєднання дозволяє використовувати інструменти стилізації, маршрутизації та анімації всередині десктопних та мобільних застосунків.

Огляд літератури. Сучасним фреймворком для створення кросплатформених застосунків із використанням звичних вебтехнологій, зокрема HTML, CSS та JavaScript, є Electron [3]. З його використанням можна перетворювати вебзастосунки, розроблені за допомогою мови програмування JavaScript, на десктопні застосунки для операційних систем Windows, MacOS та Linux. Вихідний застосунок залишається з тим самим набором програмного коду для подальшого розроблення. В основі Electron лежить комбінація двох ключових компонентів: движок Chrome V8 для виконання JavaScript та фреймворк Node.js для доступу до системних ресурсів [4]. Архітектура Electron базується на двох основних процесах: головному процесі, який керує життєвим циклом застосунку та має доступ до системних API (Application Programming Interface – прикладний програмний інтерфейс), та процесі рендерингу, який відповідає за відображення користувацького інтерфейсу. Комунікація між процесами здійснюється через вбудований механізм IPC (Inter-Process Communication) [5].

Для створення користувацьких вебінтерфейсів часто використовується декларативна бібліотека React [6]. Вона надає компонентний підхід, де інтерфейс розбивається на незалежні компоненти. React використовує віртуальний DOM (Document Object Model– об’єктна модель документа) для оптимізації рендерингу та забезпечення високої продуктивності застосунків. Також він містить однонаправлений потік даних, за рахунок чого спрощуються відстеження змін стану застосунку [7]. Програмний код JavaScript поєднується з розміткою компонентів у JSX (JavaScript XML) синтаксисі [8].

Next.js розширює можливості React [9], надаючи готовий фреймворк для створення повноцінних вебзастосунків. Next.js забезпечує серверний рендеринг SSR (Server Side Rendering), статичну генерацію сторінок SSG (Static Site Generation), автоматичний роутинг на основі файлової системи, оптимізацію зображень, вбудовану підтримку CSS модулів та API маршрути [10]. Фреймворк також надає можливості для оптимізації продуктивності, включаючи автоматичне розділення коду, попереднє завантаження сторінок та оптимізацію зображень. Next.js спрощує розроблення складних вебзастосунків, надаючи готові рішення для типових завдань та проблем [11].

Описані бібліотеки разом формують фреймворк Nextron [12], який дозволяє використовувати всі переваги Next.js, включаючи серверний рендеринг та системи маршрутизації, в контексті десктопного застосунку Electron. Тобто застосунок Next.js виконується в процесі рендерингу Electron, в той час як основна логіка застосунку та взаємодія з операційною системою здійснюється через головний процес Electron.

Створення інтерактивних графіків та діаграм для вебзастосунків може відбуватися за допомогою бібліотеки Chart.js [7]. В її основі лежить елемент HTML5 Canvas, який забезпечує високу продуктивність при відображенні графіків. Бібліотека

використовує об'єктно-орієнтований підхід, де кожен тип графіка представлений окремим класом, що успадковується від базового класу Chart [13]. Архітектура побудована за модульним принципом, дозволяючи імпортувати лише необхідні компоненти, зменшуючи загальний розмір. Підтримує щонайменше вісім основних типів діаграм, кожен з яких може бути налаштований відповідними параметрами конфігурації.

Архітектурна реалізація SIMD базується на використанні спеціалізованих векторних регістрів та набору векторних інструкцій [14]. Ці регістри мають збільшену розрядність (128, 256 або 512 біт) та можуть зберігати множину елементів даних одночасно. Процесорні інструкції SIMD оперують цими регістрами як єдиним цілим, виконуючи паралельну обробку всіх елементів за один такт процесора [15]. У сучасних процесорних архітектурах існує декілька поколінь SIMD-розширень. Для архітектури x86 це послідовність технологій MMX, SSE (Streaming SIMD Extensions) різних версій, AVX (Advanced Vector Extensions) та найновіша AVX-512. Кожне нове покоління розширень збільшувало розмір векторних регістрів та додавало нові інструкції, розширюючи можливості паралельної обробки даних.

Заголовний файл immintrin.h у Visual Studio C++ є компонентом для роботи з SIMD-інструкціями на сучасних процесорах Intel та AMD. Даний файл надає програмний інтерфейс для доступу до внутрішніх функцій процесора, які забезпечують паралельну обробку даних на апаратному рівні. Інтерфейс immintrin.h забезпечує доступ до різних наборів SIMD-інструкцій, включаючи MMX, SSE, SSE2, SSE3, SSSE3, SSE4.1, SSE4.2, AVX, AVX2 та AVX-512 [16]. Кожен з цих наборів представляє еволюційний розвиток технології SIMD, розширюючи можливості паралельної обробки даних та підвищуючи продуктивність обчислень.

Мета роботи. Метою дослідження є розроблення спеціалізованого кросплатформеного застосунку для автоматизації процесу проведення обчислювальних експериментів із порівнянням продуктивності стандартних та оптимізованих реалізацій обчислювальних алгоритмів. В якості об'єкта для дослідження було обрано задачу розв'язання СЛАР різними методами – прогонки, класним методом Гауса та методом спряжених градієнтів – для встановлення та порівняння продуктивності кожного з них. Для реалізації оптимізованої версії перелічених алгоритмів застосовано технологію SIMD. Стандартні та оптимізовані алгоритми мають бути розроблені мовою програмування C++, оскільки вона є більш продуктивною, ніж JavaScript. Десктопний застосунок повинен надавати користувацький інтерфейс для проведення експериментів, запускаючи окремо розроблені версії алгоритмів.

Дослідження спрямоване на визначення потенціалу оптимізації обчислювальних процесів за рахунок використання паралелізму на рівні даних на сучасних процесорах. Досягнення поставленої мети передбачає вирішення комплексу взаємопов'язаних завдань:

- реалізувати фундаментальні алгоритми розв'язання СЛАР із використанням засобів мови програмування C++;
- реалізувати паралельні версії зазначених алгоритмів із використанням бібліотеки immintrin.h та інструкцій AVX для оптимізації векторних операцій;
- розробити кросплатформений застосунок на Nextron для проведення обчислювальних експериментів та представлення порівняльного аналізу у вигляді таблиць і графіків;
- із використанням розробленого застосунку провести обчислювальні експерименти та проаналізувати отримані результати.

Основна частина. Класичні та оптимізовані алгоритми розв'язання СЛАР реалізовані мовою C++ як окремі складові. Для проведення обчислень вони генерують однорідні дані випадковим чином на початку своєї роботи та вимірюють час розв'язання СЛАР

класичним і оптимізованим способами, виводячи отримані результати в консольний рядок.

Розроблене програмне забезпечення представляє собою кросплатформений застосунок (рис. 1) для порівняльного аналізу швидкодії послідовних та паралельних алгоритмів розв'язання СЛАР. Основним призначенням системи є візуалізація та оцінка ефективності паралельних обчислень із використанням векторних інструкцій AVX.

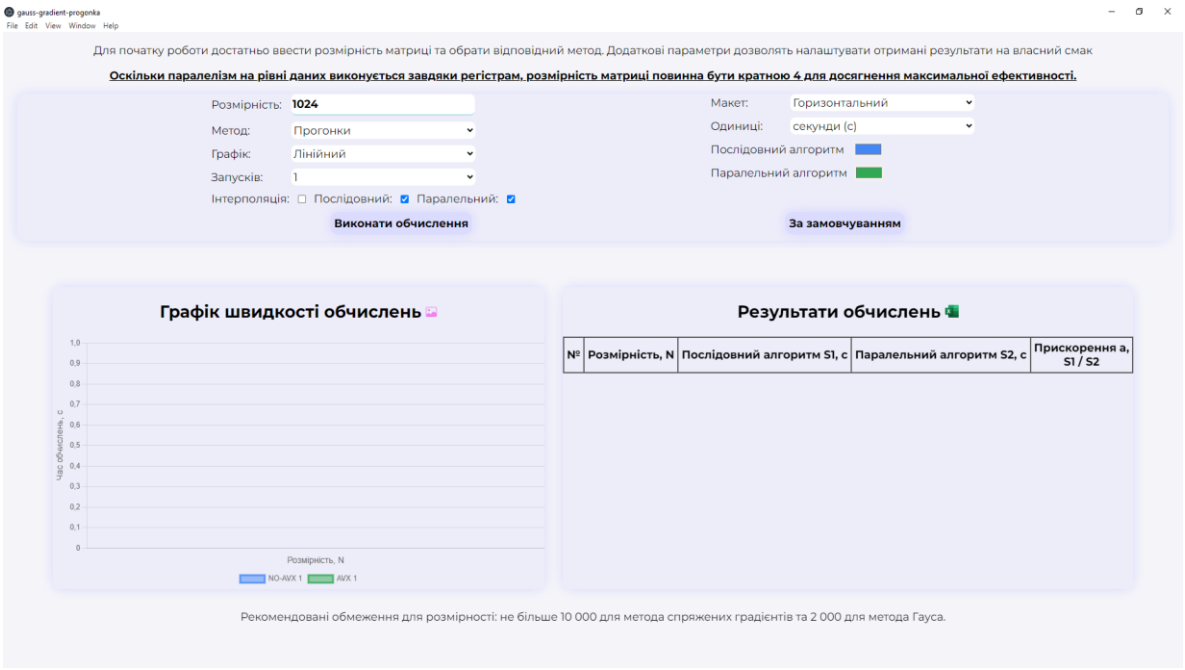


Рис. 1. Користувацький інтерфейс кросплатформеного застосунку

Архітектура програмного забезпечення побудована на основі компонентного підходу з використанням фреймворку Next.js. Центральним елементом системи є компонент Home, який керує всією логікою роботи застосунку та станом даних. Для зберігання та управління даними використовується локальний стан на основі React hooks, що дозволяє динамічно оновлювати інтерфейс при зміні параметрів обчислень.

Користувацький інтерфейс застосунку розділений на декілька функціональних блоків. У верхній частині розташована панель налаштувань, де користувач може задати основні параметри обчислень: розмірність матриці, метод розв'язання (прогонки, класичний Гауса або спряжених градієнтів), тип візуалізації результатів, кількість запусків для усереднення результатів та додаткові параметри відображення даних. Необхідною є вимога щодо кратності розмірності матриці числу 4, що обумовлено специфікою роботи векторних регістрів AVX.

Взаємодія користувача з системою описана у вигляді діаграми послідовності (рис. 2).

Процес обчислень запускається окремою функцією, яка виконує послідовний запуск алгоритмів, розроблених на C++, для різних розмірностей матриці. Система автоматично розбиває діапазон розмірностей на фіксовані кроки для побудови графіків залежності часу обчислень від розміру задачі. Результати кожного запуску зберігаються окремо для послідовного та паралельного алгоритмів з метою подальшого проведення порівняльного аналізу.

Візуалізація результатів реалізована за допомогою компонента Chart, який надає можливість відображення даних у різних форматах: лінійний графік, гістограма, радарна діаграма, бульбашкова діаграма, полярна діаграма тощо. Для кожного типу візуалізації передбачені додаткові налаштування, такі як інтерполяція даних та

кольорове оформлення графіків. Система також включає табличне представлення результатів, де для кожної розмірності задачі відображаються час виконання послідовного та паралельного алгоритмів, а також досягнуте їх пришвидшення. Передбачена можливість експорту результатів у форматі Excel та збереження графіків у вигляді PNG-зображень для подальшого аналізу. За необхідності можна динамічно змінювати одиниці вимірювання часу (секунди або мілісекунди), налаштовувати макет відображення результатів (горизонтальний або вертикальний), та керувати відображенням окремих компонентів візуалізації. Система також надає рекомендації щодо обмежень на розмірність задачі для різних методів розв'язання, допомагаючи користувачу обрати оптимальні параметри для конкретного випадку.

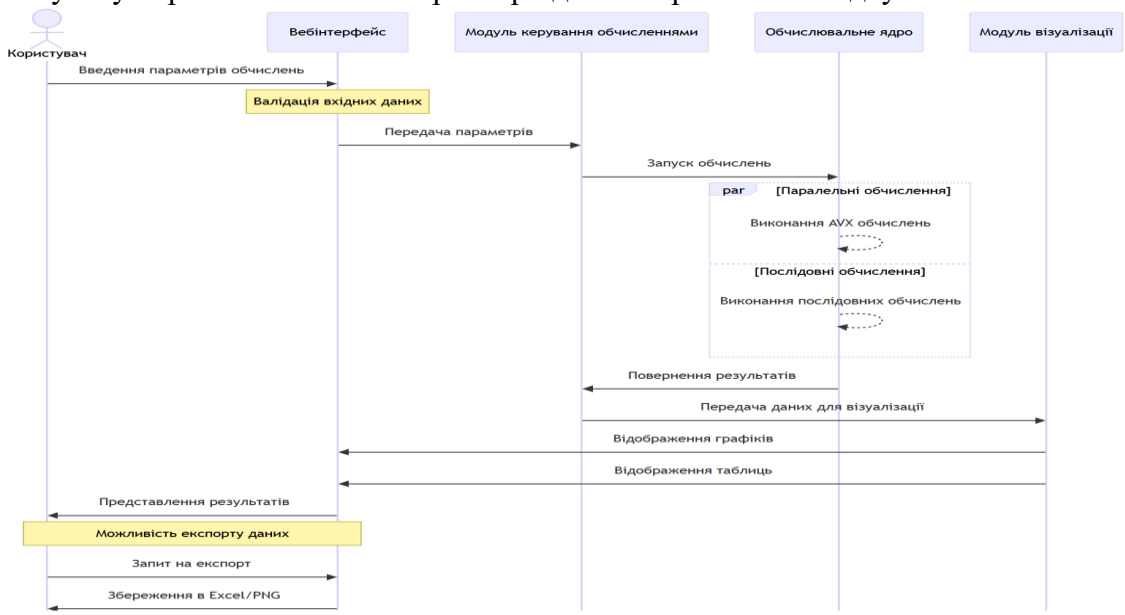


Рис. 2. Діаграма послідовності взаємодії користувача з системою

Обробка станів завантаження та помилок забезпечує стабільну роботу системи навіть при тривалих обчисленнях. Програмне забезпечення автоматично блокує елементи керування під час виконання обчислень та надає візуальний зворотний зв'язок про стан процесу. Архітектура програмної системи розділена на декілька ключових модулів. Модуль вебінтерфейсу відповідає за представлення користувацького інтерфейсу та взаємодію з користувачем. Він реалізований за допомогою React-компонентів та включає в себе форми введення параметрів, елементи керування та контейнери для відображення результатів. Модуль керування обчисленнями є центральним компонентом системи, який координує роботу інших модулів. Він обробляє користувацькі введення, виконує валідацію параметрів, керує процесом обчислень та розподіляє результати між іншими компонентами. У цьому модулі реалізована основна бізнес-логіка програми, включаючи функцію для запуску обчислень та обробку отриманих результатів. Модуль обчислювального ядра відповідає за безпосереднє виконання математичних розрахунків. Він розділений на два підмодулі: модуль послідовних алгоритмів та модуль паралельних обчислень з використанням AVX. Кожен з цих підмодулів виконує відповідні алгоритми розв'язання СЛАР. Модуль візуалізації відповідає за графічне представлення результатів обчислень. Підмодуль відображення графіків використовує компонент Chart для побудови різних типів діаграм, а підмодуль табличного представлення формує структуровані таблиці результатів. Модуль експорту даних забезпечує можливість збереження результатів у різних форматах. Він взаємодіє з компонентами візуалізації для експорту графіків у форматі PNG та таблиць у форматі Excel, використовуючи бібліотеки XLSX та html2canvas для реалізації функціональності

експорту. Модуль управління станом забезпечує централізоване зберігання та управління даними програми. Він реалізований з використанням React hooks (хук useState) та відповідає за синхронізацію стану між різними компонентами системи. Цей модуль зберігає такі дані як результати обчислень, налаштування візуалізації, параметри користувацького інтерфейсу тощо. Модуль конфігурації містить налаштування системи, включаючи параметри за замовчуванням, обмеження на розмірності матриць для різних методів, колірні схеми для візуалізації та інші константи програми. Він також відповідає за валідацію користувацьких налаштувань. Взаємодія між модулями організована за принципом слабкого зв'язування, що забезпечує можливість незалежного розвитку та модифікації кожного з них. Система використовує події та зворотні виклики для комунікації між модулями, підтримуючи асинхронну природу застосунку.

Результат та обговорення. В результаті розроблення було реалізовано кросплатформений застосунок для автоматизованого проведення обчислювальних експериментів. Дослідження ефективності паралельних обчислень проводилося на основі порівняльного аналізу часу виконання послідовних та паралельних реалізацій трьох різних методів розв'язання СЛАР. Паралельна реалізація базується на використанні технології SIMD за допомогою набору інструкцій AVX. Для проведення експериментів використовувалася розроблена програмна система. Результати вимірювань часу виконання алгоритмів представлені у вигляді графіків залежності часу обчислень від розмірності задачі. Для забезпечення точності вимірювань система підтримує представлення результатів як у секундах, так і в мілісекундах, дозволяючи детально аналізувати продуктивність на різних масштабах даних. Головним показником ефективності паралельної реалізації є коефіцієнт прискорення. Експериментальні дані демонструють стабільне пришвидшення для всіх досліджуваних методів, при цьому спостерігається нелінійна залежність коефіцієнта прискорення від розмірності задачі. На рис. 3 наведено результати одного з обчислювальних експериментів. Для тестування був використаний метод спряжених градієнтів при максимальній розмірності матриці – 2048. Всі три запуски надали приблизно однаковий результат. Досягнуте пришвидшення обчислень знаходиться в межах 2.67–5.81 в залежності від розмірності матриці. Для меншої розмірності матриці фіксується менше пришвидшення через накладні витрати на векторизацію.

Висновки. Розроблений програмний комплекс дозволяє автоматизувати проведення експериментів з порівняльного аналізу обчислювальних методів із використанням технології SIMD. Результати засвідчують, що використання SIMD-інструкцій сприяє підвищенню продуктивності обчислень, зокрема для задач, які включають обробку великих об'ємів даних. Кросплатформений застосунок може використовуватись для проведення експериментів та візуалізації отриманих результатів будь-яких окремо розроблених обчислювальних алгоритмів. Експерименти, проведені в межах дослідження, виявили суттєве пришвидшення виконання алгоритмів розв'язання СЛАР, таких як методи прогонки, класичний метод Гауса та метод спряжених градієнтів. Залежно від розмірності задачі та методу, коефіцієнт прискорення варіювався в межах від 2.67 до 5.81 разів. Зокрема, для малих розмірностей матриць ефект пришвидшення обмежений через накладні витрати на векторизацію, однак зі збільшенням об'ємів даних ефективність SIMD-інструкцій стає більш помітною. Отримані результати пришвидшення обчислень кореспондуються з результатами досліджень авторів [17].

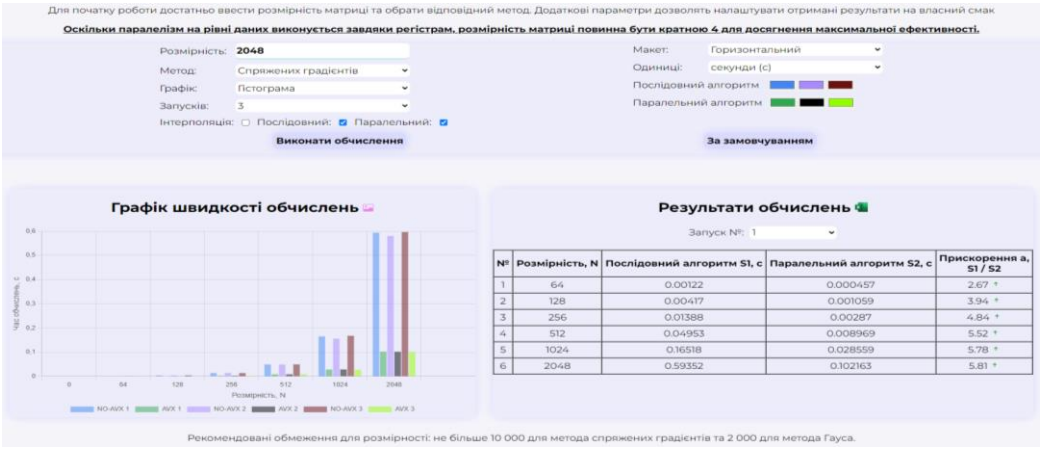


Рис. 3. Демонстрація результатів обчислювального експерименту

Описана архітектура програмного забезпечення базується на модульному підході та забезпечує гнучкість і масштабованість. Використання компонентного підходу з React та Next.js дозволяє підтримувати асинхронність та інтерактивність для сучасних систем, орієнтованих на обробку великих об’ємів даних. Модулі, відповідальні за візуалізацію, дозволяють представляти результати у вигляді графіків та таблиць, а також експортувати їх для подальшого аналізу, що розширює можливості використання системи в дослідницьких цілях. Дослідження також підтверджує доцільність і ефективність застосування SIMD-інструкцій для розв’язання обчислювальних задач, демонструючи переваги паралельної на рівні даних обробки у порівнянні зі стандартними підходами. Отримані результати можуть бути використані для розроблення нових комп’ютерних моделей та систем, орієнтованих на високопродуктивні обчислення.

Список літератури

1. Zhang W., Yan Z., Lin Y., Zhao C., Peng L. A High Throughput B+tree for SIMD Architectures. *IEEE Trans. Parallel Distrib. Syst.* 2020. V. 31. No. 3. P. 707–720. URL: https://doi.org/10.1109/tpds.2019.2942918
2. Naganawa Y., Kamei H., Kanetaka Y., Nogami H., Maeda Y., Fukushima N. SIMD-Constrained Lookup Table for Accelerating Variable-Weighted Convolution on x86/64 CPUs. *IEEE Access*. 2024. V. 12. P. 15800-15819. URL: https://doi.org/10.1109/access.2024.3354720
3. Kredpattanakul K., Limpiyakorn Y. Transforming JavaScript-Based Web Application to Cross-Platform Desktop with Electron. *Information Science and Applications 2018*. Singapore. 2018. P. 571–579. URL: https://doi.org/10.1007/978-981-13-1056-0_56
4. Srinivasa Rao M., Sandhya P., Sambana B., Mishra P. Application for Mood Detection of Students Using TensorFlow and Electron JS. *Springer Proceedings in Mathematics & Statistics*. Cham, 2023. P. 235–243. URL: https://doi.org/10.1007/978-3-031-15175-0_19
5. Electron Documentation. URL: https://www.electronjs.org/docs/latest/
6. Shevtsiv N.A., Striuk A.M. Cross platform development vs native development. *CEUR Workshop Proceedings*. 2021. V. 2832. P. 75–83. URL: https://doi.org/10.31812/123456789/4428
7. Bernatska N., Typilo I., Dzhumelia E. React Library: A Case Study on the Effective Instruments of the Design of an Environmental Monitoring System. *2023 IEEE 18th Int. Conf. Comput. Sci. Inf. Technol. (CSIT)*. 2023. P. 1–4. URL: https://doi.org/10.1109/csit61576.2023.10324124
8. React Documentation. URL: https://legacy.reactjs.org/docs/getting-started.html
9. Patel V. Analyzing the Impact of Next.JS on Site Performance and SEO. *International Journal of Computer Applications Technology and Research*. 2023. V. 12. P. 24–27. URL: https://doi.org/10.7753/IJCATR1210.1004

10. Jartarghar H.A., Rao S. G., Ashok K.A.R, Sharvani G.S., Dalali S. React Apps with Server-Side Rendering: Next.js. *J. Telecommunication, Electron. Comput. Eng. (JTEC)*. 2022. V. 14, no. 4. P. 25–29. URL: https://doi.org/10.54554/jtec.2022.14.04.005
11. Next.js Documentation. URL: https://nextjs.org/docs
12. Nextron Github. URL: https://github.com/saltyshiomix/nextron
13. Chart.js Documentation. URL: https://www.chartjs.org/docs/latest
14. Mustafa D., Alkhasawneh R., Obeidat F., Shatnawi A.S. MIMD Programs Execution Support on SIMD Machines: A Holistic Survey. *IEEE Access*. 2024. V. 12, P. 34354–34377. URL: https://doi.org/10.1109/access.2024.3372990
15. Intel Architecture Instruction Set Extensions Programming Reference. URL: https://www.intel.com/content/dam/develop/external/us/en/documents/319433-024-697869.pdf
16. Intel Intrinsics Guide. URL: https://www.intel.com/content/www/us/en/docs/intrinsics-guide/index.html
17. Zhulkovskyi O.O., Zhulkovska I.I., Vokhmianin H.Ya., Firsov O.D., Riabovolenko V.A. Research of progressive tools of parallel computations with the use of SIMD architecture. *Inform. math. methods simul*. 2023. V. 13, no. 3-4. P. 228–235 URL: https://doi.org/10.15276/imms.v13.no3-4.228

CROSS-PLATFORM SYSTEM FOR ANALYZING THE EFFICIENCY OF PARALLEL COMPUTING ALGORITHMS

O. O. Zhulkovskyi[1], H. Ya. Vokhmianin[1], I. I. Zhulkovska[2],
Yu. V. Ulianovska[2], V. A. Riabovolenko[2]

[1]Dniprovsky State Technical University
2, Dniprobudivska Ave., Kamianske city, 51918, Ukraine
[2]University of Customs and Finance
2/4, Volodymyr Vernadskyi Ave., Dnipro, 49000, Ukraine
Email: olalzh@ukr.net

The paper presents the results of the development and research of a specialized system for automated comparative analysis of the efficiency of computational algorithms, in particular, using SIMD technology. The main goal of the study is to create cross-platform software for conducting computational experiments that allows comparing the performance of standard and optimized implementations of algorithms for solving systems of linear algebraic equations. The study focuses on approaches to the development and comparison of the performance of fundamental methods for solving equations, in particular, Gauss and conjugate gradient methods. For the studied methods, two versions of the implementation were developed – standard and optimized using SIMD instructions. The software implementation of the algorithms was performed by the productive tools of Microsoft Visual Studio C++ using the standard specialized library immintrin.h and the AVX processor instruction set. The developed software system is based on a modern technology stack, in particular, it uses the Nextron framework, which combines the capabilities of Electron for creating cross-platform applications and Next.js for building an interactive user interface. The architecture of the system provides modularity, scalability, and usability through the use of the React component approach. The paper presents the results of an experimental study of the effectiveness of using SIMD technology to speed up computational algorithms. The experiments demonstrate a significant increase in the performance of parallel computing algorithms in the range of 2.67–5.81 times, depending on the dimensionality of the SLAE and the chosen computational method. It was found that the efficiency of SIMD optimization increases proportionally to the increase in the amount of input data, while for small dimensions the vectorization overhead can limit the overall performance. The results of the study confirm the potential of using SIMD technology to optimize computational algorithms and demonstrate the practical value of the developed software for automating the process of conducting and analyzing computational experiments. The results obtained can be used in the development of other optimized algorithms and software systems, in particular for computer modeling, focused on high-performance computing.
Keywords: cross-platform application, SIMD, data-level parallelism, computing acceleration, immintrin.h, Electron, Next.js, Nextron.

DOI 10.15276/imms.v14.no4.358
УДК 621.438

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 358-365

ВДОСКОНАЛЕННЯ СИСТЕМИ АВТОМАТИЧНОГО УПРАВЛІННЯ КОГЕНЕРАЦІЙНОЮ ЕНЕРГЕТИЧНОЮ УСТАНОВКОЮ НА БАЗІ ГТУ

О. Є. Мішкой, О.С. Тарахтій

Національний університет «Одеська політехніка»
1, Шевченка пр., м.Одеса, 65044, Україна
Email:mishkoy2002@gmail.com

Розглянуто питання підвищення якості управління потужністю газотурбінної установки, яка є тепловим двигуном когенераційної енергетичної установки. Актуальність теми зумовлена складною економічною ситуацією в енергетичній галузі України, високими цінами на імпортні енергоресурси та необхідністю раціонального використання місцевих і альтернативних видів палива. У роботі акцентовано на важливості підвищення енергоефективності ГТУ, зокрема за допомогою автоматизації управління, що дозволяє забезпечити стабільну роботу установки за різних зовнішніх умов і типів палива. Головною метою роботи була розробка структури системи автоматичного управління потужністю ГТУ, яка враховує зовнішні збурення, такі як зміна електричного навантаження, теплоти згоряння палива та температури зовнішнього повітря. Запропоновано використання додаткового сигналу за похідною від температури робочого тіла перед турбіною, що сприяє значному покращенню показників перехідних процесів. Наведено структурну схему автоматичного управління, проведено моделювання перехідних процесів регулювання частоти обертання ротора турбіни, як найголовнішого регульованого параметру та проведення порівняльний аналіз якості регулювання при введенні додаткового сигналу за похідною по температурі газів на вході до турбіни. Результати моделювання показали, що використання додаткового сигналу за похідною підвищує стійкість системи, зменшує динамічне відхилення параметрів до 46,8% та покращує затухання коливань на 16%. Також впровадження автоматизованого управління із сигналом за похідною зменшує відхилення температури газів на виході з камери згоряння, що дозволяє знизити термічне напруження на елементи проточної частини ГТУ.
Ключові слова: газотурбінні установки, когенерація, автоматичне управління, перехідні режими, моделювання, зовнішні збурення, ГТУ.

Вступ. Сучасний стан енергетики України, руйнування енергогенеруючих потужностей, а також умови глобальної енергетичної та екологічної політики вимагають негайного впровадження енергоефективних технологій і створення розподіленої системи генерації електричної енергії. Таке рішення дозволить зробити енергосистему нашої країни менш уразливою до ракетних обстрілів і створити окремі райони держави енергонезалежними [1,2]. Когенерація є одним із таких рішень, яке може дозволити розподілити об’єкти енергогенерації за регіонами нашої держави, в яких є найбільший дефіцит генеруючих поутжностей. Використання когенераційних технологій також дозволяє суттєво підвищити загальну ефективність використання палива порівняно з традиційними роздільними системами виробництва електричної енергії. Для підвищення ККД даного типу установок, тепло продуктів згоряння від теплового двигуна, використовують для вироблення додаткової теплової енергії, цей підхід є одним з ключових напрямків енергозбереження.

Мета роботи. Метою є вдосконалення автоматизованої системи регулювання газотурбінною установкою введенням в регулятор додаткового імпульсу за похідною по температурі газів після камери згоряння і проведення аналізу якості отриманих перехідних процесів регулювання.

Модель когенераційної установки на базі ГТУ. В якості теплових двигунів в системах когенерації часто використовуються газотурбінні установки. Це обумовлено, головним

чином, їхніми маневровими характеристиками і малим часом запуску, а також достатньо великою одиничною потужністю окремої установки [3–5].

Когенераційні установки представляють собою системи, які призначені для одночасного вироблення теплової та електричної енергії з використанням одного джерела палива. В основі їх роботи лежить принцип комбінованого виробництва, який дозволяє значно підвищити коефіцієнт корисної дії (ККД) системи за рахунок утилізації тепла, що виділяється під час виробництва електроенергії.

Когенераційна установка складається з наступних ключових елементів: 1.Первинний двигун – це джерело механічної енергії, який може бути газовою, паровою турбіною або двигуном внутрішнього згоряння. 2. Електрогенератор – перетворює механічну енергію первинного двигуна на електричну енергію. 3. Система утилізації тепла – ключова частина когенераційної установки, яка підвищує загальну ефективність системи. До неї входять теплообмінні апарати, такі як котли-утилізатори та теплообмінники, які забезпечують утилізацію тепла, що виділяється первинним двигуном. 4. Система управління та контролю – відповідає за моніторинг та регулювання параметрів установки, таких як температура, тиск, витрата палива та повітря, частота обертання генератора тощо.

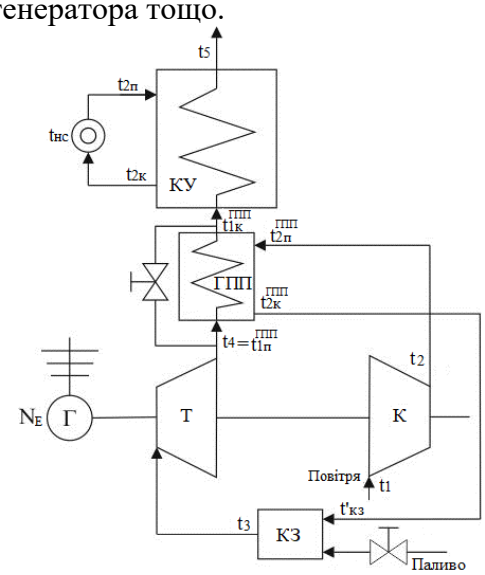


Рис. 1. Структурна схема когенераційної установки на базі ГТУ з утилізацією теплоти продуктів згоряння

Потенційними об'єктами для застосування когенерації є промислові виробництва, заводи, підприємницька сфера (пекарні, хімчистки тощо) лікарні, готелі, торгові центри, адміністративні центри, ферми, об'єкти житлової сфери – житлові будинки та приватні будинки, громадські установи: лікарні, курортні та лікувальні заклади, басейни, спортивні центри, казарми тощо, власні потреби газоперекачувальних станцій, компресорних станцій, котелень тощо.

Когенераційні установки можуть використовуються не тільки як незалежні альтернативні міні ТЕЦ, а також в якості резервних, допоміжних джерел тепло- та електроенергії. Вироблена електрична енергія, може бути підключена до спільної розподільної мережі або використана у власній мережі. Аналогічним чином, тепло, яке утворюється під час спалювання палива в ГТУ, може бути приєднана до централізованих тепломереж або використана в якість води для опалення та виробництва гарячої води для побутових потреб.

Для моделювання динамічних властивостей і перехідних процесів когенераційної установки була використана математична модель наведена у [6]. Дана модель представляє собою систему з 6 диференційних рівнянь. Перші три рівняння описують динамічні властивості газоповітряного підігрівача (ГПП), а наступні три

диференційні рівняння описують динаміку ротора ГТУ, зміни тиску у газових об'ємах і температури газів після камери згоряння.

$$
\begin{aligned}
&\mathrm{C}^{\text{ГПП}}\frac{d\Delta t_{2\text{к}}^{\text{ГПП}}}{dt}+\Delta t_{2\text{к}}^{\text{ГПП}}=\mathrm{c}_1^{\text{ГПП}}\Delta t_2+\mathrm{c}_2^{\text{ГПП}}\Delta t_{\text{ст}}^{\text{ГПП}};\\
&H^{\text{ГПП}}\frac{d\Delta t_{1\text{к}}^{\text{ГПП}}}{dt}+\Delta t_{1\text{к}}^{\text{ГПП}}=h_1^{\text{ГПП}}\Delta t_{1\text{н}}^{\text{ГПП}}-h_2^{\text{ГПП}}\Delta t_{\text{ст}}^{\text{ГПП}};\\
&R^{\text{ГПП}}\frac{d\Delta t_{\text{ст}}^{\text{ГПП}}}{dt}+\Delta t_{\text{ст}}^{\text{ГПП}}=r_1^{\text{ГПП}}\Delta t_{1\text{н}}^{\text{ГПП}}+r_2^{\text{ГПП}}\Delta t_{1\text{к}}^{\text{ГПП}}-r_3^{\text{ГПП}}\Delta t_2-r_4^{\text{ГПП}}\Delta t_{2\text{к}}^{\text{ГПП}};\\
&B\frac{\partial\Delta\omega}{\partial t}+\Delta\omega=b_1\Delta p_3+b_2\Delta t_3+b_3\Delta p_4-b_4\Delta p_2-b_5\Delta N_{\text{E}};\\
&\mathrm{T_p}\frac{\partial\Delta p}{\partial t}+\Delta p=T_{\text{T}}\frac{d\Delta T_3}{dt}-k_{\text{T}}\Delta T_3+k_m\Delta m_{\text{п}}+k_\omega\Delta\omega;\\
&\mathrm{A}\frac{\partial\Delta t_3}{\partial t}+\Delta t_3=\mathrm{a}_1\Delta\omega+\mathrm{a}_2\Delta m_{\text{п}}+\mathrm{a}_3\Delta t_2.
\end{aligned}
\tag{1}
$$

Структурна схема імітаційної моделі, побудована на базі вищевказаної системи диференційних рівнянь, наведена на рис. 2.

На схемі прийняті наступні позначення: *GVP* – модель газоповітряного підігрівача; *Gas volumes* – модель зміни тиску у газових об'ємах ГТУ; *Rotor of turbine* – модель зміни частоти обертання ротора ГТУ; *KC* – камера згоряння. Блоки Compressor і *Turbine* моделюють процеси адіабатичного стискання повітря у компресорі та адіабатичного розширення газів в турбіні.

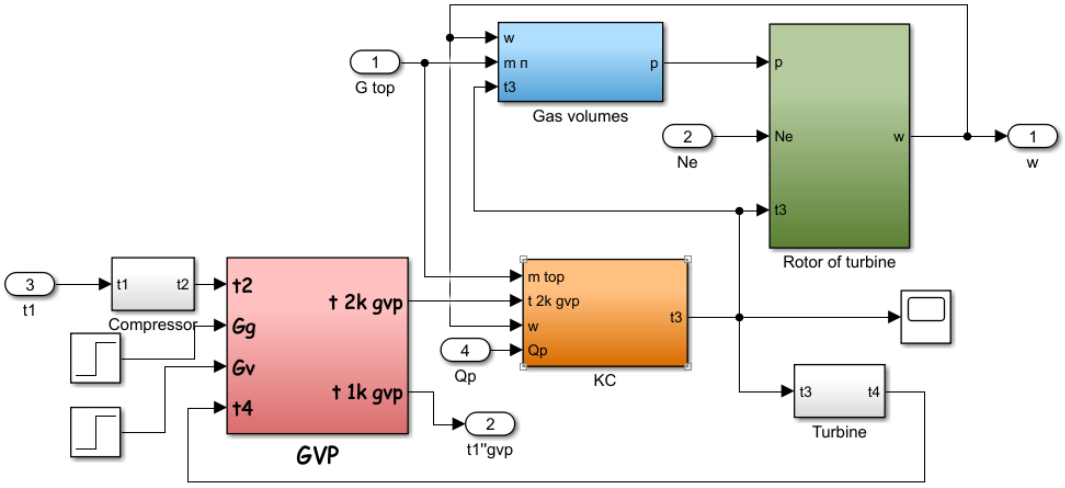


Рис. 2. Структурна схема моделі когенераційної установки на базі ГТУ

Наведена модель представляє собою систему лінеаризованих диференційних рівнянь із зосередженими параметрами і описує динаміку когенераційної установки в умовах малих відхилень від стаціонарного режиму.

Вдосконалення системи автоматизованого управління когенераційною установкою. Як відомо ККД стаціонарних ГТУ, що працюють за циклом Брайтона, невеликий, і може становити від 24 до 32 % [7,8]. Для збільшення ККД цих установок застосовують утилізацію теплоти відхідних газів у котлах-утилізаторах, для вироблення теплової енергії та регенерацію теплоти за допомогою газо-повітряних підігрівачів, для підігріву повітря після компресора [7,8].

Однак установки такого типу окрім теплової енергії виробляють також і електричну енергію. У вітчизняній енергосистемі висуваються достатньо жорсткі вимоги до стабілізації частоти електричної енергії [11], яка надходить до енергосистеми: відхилення частоти не повинне перевищувати ± 0,2 Гц. Цей факт підвищує вимоги до якості автоматичного регулювання когенераційною установкою з точки зору стабілізації частоти виробленої електричної енергії. За основу була взята система автоматичного

управління газотурбінною установкою, наведена у [8,9]. Умови роботи енергогенеруючого обладнання у частих перехідних режимах, які обумовлені складним станом вітчизняної енергосистеми та складністю розподілення навантаження в енергосистемі підтверджують актуальність задачі підвищення якості та сталості автоматичної системи регулювання частоти. Нижче наведена структурна схема автоматизованої системи управління потужністю газотурбінної установки, яка була вдосконалена введенням додаткового сигналу за похідною по температурі продуктів згоряння на виході з камери згоряння (рис. 3).

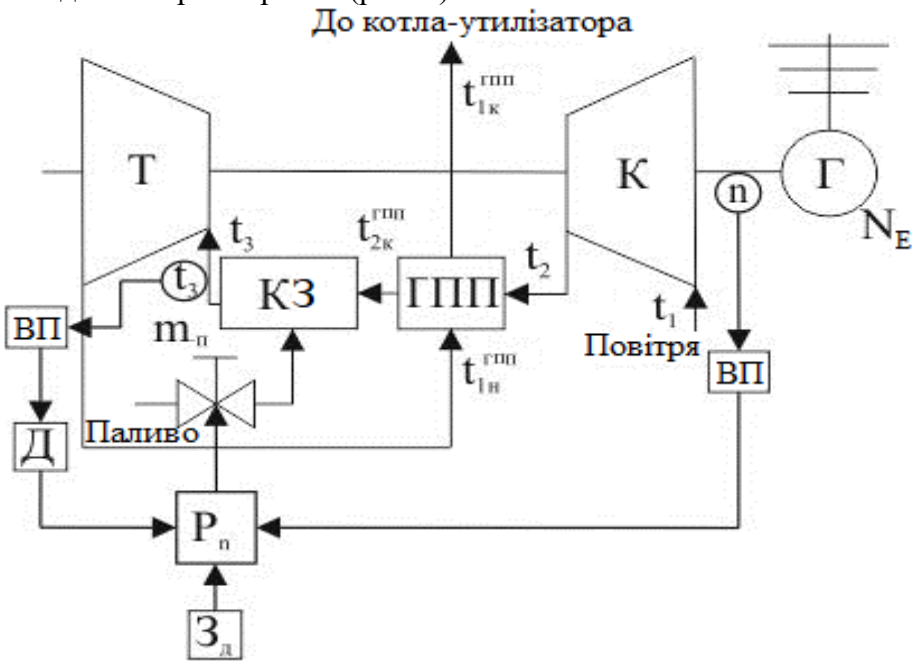


Рис. 3. Структурна схема автоматизованої системи управління ГТУ когенераційної установки

Основними елементами на наведеній схемі є: камера згоряння (КЗ) до якої надходить повітря, що нагнітається компресором (К) і паливо через регулюючий орган ($m_п$). А також, турбіна (Т), на ротор якої діють рухомі сили газів, які виходять з камери згоряння і сили опору з боку електричного генератора (Г) і компресора. Повітря після компресора нагрівається у газо-повітряному підігрівачі (ГПП) від теплоти газів, які виходять після турбіни. За рахунок цього знижується витрата палива на підігрів повітря у камері згоряння.

Основними регулюючими параметрами в ГТУ є: частота обертання ротора газової турбіни ω і температура газів на виході з камери згоряння t_3 [10]. Частота обертання ротора ГТУ однозначно пов'язана із частотою обертання валу генератора, генеруючим електричну енергію і, відповідно частотою електричного струму, що надходить до енергосистеми. Температура газів на виході камери згоряння у сучасних ГТУ складає близько 2000 °С і визначає термічне напруження елементів проточної частини газової турбіни, а саме робочих лопаток. Навіть короткочасне підвищення температури газів може суттєво вплинути на довговічність роботи або привести до виходу з ладу елементів проточної частини.

Введення додаткового сигналу за похідною від значення температури газів після камери згоряння може дозволити підвищити якість регулювання температури, роблячи регулюючий вплив, пропорційним швидкості зміни температури продуктів згоряння. А також може підвищитися швидкість регулювання, оскільки регулятор буде змінювати положення регулюючого органу до того моменту, як відбудеться зміна частоти обертання ротора газової турбіни.

Моделювання і аналіз динамічних характеристик когенераційної установки. При

проведенні моделювання перехідних процесів наносилося збурення зміною електричного навантаження, величина збурення складала 10% від номінального навантаження газотурбінної установки.

Нижче, на рис.4 і рис.5, наведені перехідні процеси регулювання, отримані при застосуванні автоматичної системи регулювання без введення додаткового сигналу за похідною за температурою газів після камери згоряння.

Перехідний процес регулювання частоти обертання ротора ГТУ без введення додаткового сигналу за похідною наведений на рис.4

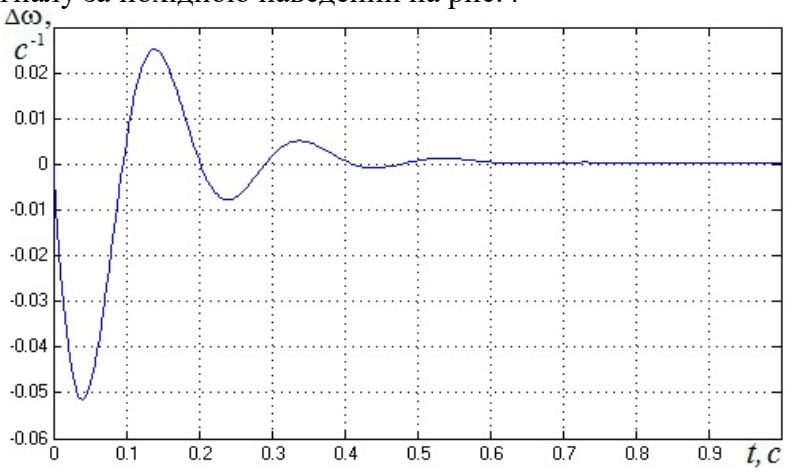


Рис. 4. Перехідний процес регулювання по каналу зміна електричного навантаження →частота обертання ротора ГТУ $\Delta N_{\text{E}} \rightarrow \omega$ (збурення $\Delta N_{\text{E}} = 10\%$)

Із графіку перехідного процесу, представленого на рис. 3 видно, що збільшення електричного навантаження ΔN_{E} на 10% призводить до відхилення частоти обертання ротора ГТУ на -0,052 Гц (-0,104 %), яке не виходить за межі допустимих відхилень, що регламентуються стандартом і складають ± 0,2 Гц. Ступінь загасання перехідного процесу при цьому становить $\psi^{N_E \rightarrow \omega} = 0{,}85$.

Аналіз моделювання динаміки системи автоматичного керування когенераційною установкою без введення додаткового сигналу за похідною по температурі газів показав, що зміна електричного навантаження генератора ΔN_{E} також викликає істотне відхилення температури газів на виході з камери згоряння (рис. 5).

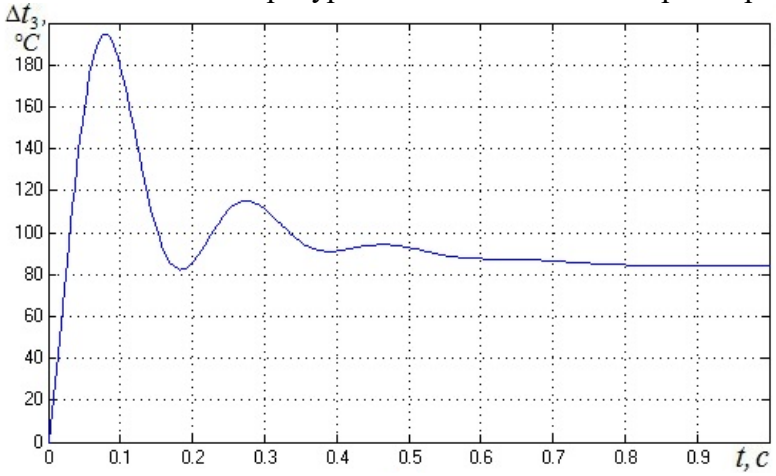


Рис. 5. Перехідний процес регулювання по каналу зміна електричного навантаження → температура газів після КЗ $\Delta N_{\text{E}} \rightarrow t_3$ (збурення $\Delta N_{\text{E}} = 10\%$)

Збільшення електричного навантаження на 10 % призводить до збільшення температури газів на виході КЗ на 16,3 % (195 °С), що небажано, оскільки збільшується термічна напруга елементів проточної частини газової турбіни. Ступінь загасання перехідного процесу, представленого складає $\psi^{N_E \rightarrow t_3} = 0{,}71$. З метою нейтралізації вказаного негативного факту, було прийнято рішення в контур регулювання

електричним навантаженням генератора додати сигнал за похідною від зміни температури газів на виході з камери згоряння (рис. 3). Для отримання похідної застосовано реальну диференційну динамічну ланку, отриману як результат алгебраїчної суми передавальних функцій пропорційної та інерційної ланок [12].Результати моделювання перехідних процесів зміни частоти обертання ротора ГТУ при збуренні зміною електричного навантаження із використанням додаткового сигналу за похідною і без нього представлені на рис. 6.

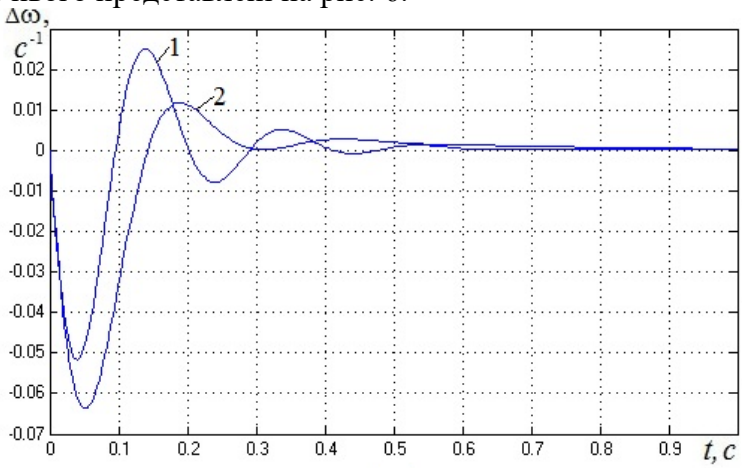


Рис. 6 Перехідні процеси регулювання частоти обертання ротора при використанні ПІ-закону і додаткового сигналу за похідною (ПІ+Д): 1 – $\Delta N_{\text{E}} = 10\%$ (ПІ); 2 – $\Delta N_{\text{E}} = 10\%$ (ПІ+Д)

Аналіз отриманих перехідних процесів показує якісне поліпшення показників перехідних процесів: збільшення ступеня загасання до 0,95 і зменшення другого динамічного відхилення, що свідчить про збільшення запасу стійкості системи. Незначне збільшення першого динамічного відхилення пояснюється тим, що збурення зміною електричного навантаження практично миттєво впливає на частоту обертання ротора, проте на другому відхиленні вже позначається вплив сигналу за похідною, і його зниження становить 46,8 %. Перехідні процеси зміни температури газів на виході з камери згоряння представлені на рис. 7

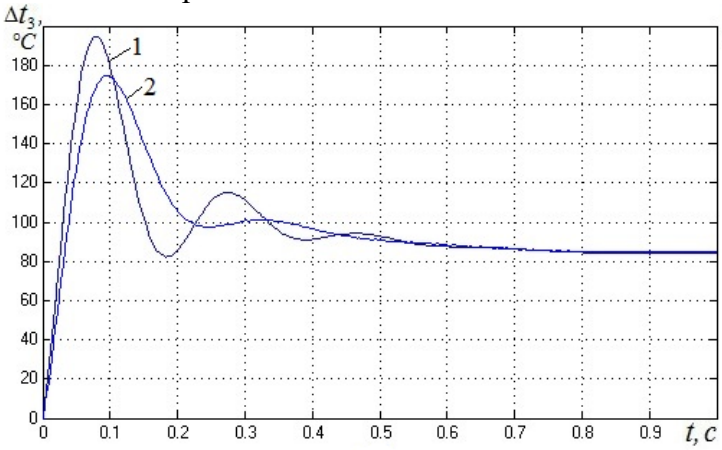


Рис. 7 Перехідні процеси регулювання температури газів на виході КЗ при використанні ПІ-закону і сигналу за похідною (ПІ+Д): 1 – $\Delta N_{\text{E}} = 10\%$ (ПІ); 2 – $\Delta N_{\text{E}} = 10\%$ (ПІ+Д)

З наведених графіків також видно, що з введенням сигналу по похідній від температури газів після КЗ відбувається покращення якості перехідного процесу і, що саме головне, знижується відхилення температури газів після камери згоряння зі 195 °С до 168°С. Цей факт дуже позитивно впливає на елементи проточної частини газової турбіни, знижуючи їхнє термічне напруження і дозволяє подовжити їх строк служби.

Висновки. Проведене дослідження підтвердило важливість вдосконалення системи

автоматичного управління когенераційними установками на базі газотурбінних двигунів для забезпечення їх стабільної та енергоефективної роботи в умовах змінних зовнішніх впливів. Результати моделювання показали, що введення додаткового сигналу за похідною температури газів після камери згоряння суттєво покращує динамічні характеристики системи. Зокрема, було досягнуто зменшення динамічних відхилень температури газів на виході з камери згоряння на 13,8%, а також підвищення ступеня загасання коливань до 95%, що знижує термічне напруження елементів газотурбінного двигуна та сприяє продовженню їх строку служби.

Впровадження такого підходу до автоматизованого управління дає змогу забезпечити високу точність регулювання частоти обертання ротора та стабільність виробленої електричної енергії, що є критично важливим для когенераційних установок у сучасній енергетичній системі. Позитивні результати моделювання свідчать про перспективність подальшого використання запропонованого методу для поліпшення роботи ГТУ, особливо в умовах частих змін навантаження та варіативності зовнішніх параметрів, таких як температура та якість палива.

Отримані результати відкривають нові можливості для підвищення енергоефективності когенераційних установок за рахунок удосконалення систем автоматичного управління. Це, у свою чергу, сприятиме більш раціональному використанню енергоресурсів і зміцненню енергетичної незалежності країни, особливо в умовах зростаючих вимог до стабільності та екологічності енергетичних систем.

Загалом, впровадження сучасних методів регулювання та автоматизації у роботу когенераційних установок на базі ГТУ є важливим кроком до розвитку стійкої та енергоефективної енергетичної інфраструктури.

Список літератури

1. Криволап К. Українська енергосистема 2023-2024: проблеми, виклики та перспективи. URL: https://rubryka.com/blog/ukrayinska-energosystema
2. Розвиток когенераційних станцій в Україні як відповідь на енергетичні виклики війни. URL: https://clearenergy.ua/rozvytok-kogeneratsijnyh-stantsij-v-ukrayini-yak-vidpovid-na-energetychni-vyklyky-vijny/
3. Цьогоріч Україна значно збільшить кількість газотурбінних та біогазових установок: Українське національне інформаційне агентство «Укрінформ». URL: https://www.ukrinform.ua/rubric-economy/3685024-cogoric-ukraina-znacno-zbilsit-kilkist-gazoturbinnih-ta-biogazovih-ustanovok-ekspert.html
4. Проблеми встановлення газотурбінних установок – синхронізація з мережею, модернізація розподільчого обладнання, URL: https://interfax.com.ua/news/economic/1033532-amp.html
5. Колотило Д.В. Лесько В.О., Гресько А.О., Кравець В.В. Газотурбінні електростанції в Об'єднаній енергетичній системі України в умовах інтеграції до Європейської мережі операторів системи передачі енергії. *LI Науково-технічна конференція факультету електроенергетики та електромеханіки*: матеріали конференції. Вінниця : ВНТУ, 2022. 3 с.
6. Тодорцев Ю.К., Ларіонова О.С. Математична модель динаміки когенераційної енергетичної установки. *Автоматика–2011*: матеріали XVIII Міжнар. конференції з автоматичного управління. Львів, 2011. С. 50 – 53.
7. Ганжа О.М., Марченко М.А. Удосконалення стаціонарної газотурбінної установки вибором раціональних параметрів регенератора-повітропідігрівача. *Вісник НТУ ХПІ: Енергетичні та теплотехнічні процеси та устаткування.* 2012. №7. С. 124 – 128.
8. Tarakhtii O., Bundyuk A. The research of the energy characteristics of cogeneration power plant in a changing fuel quality. *Automation of technological and business processes*. 2016. №8/1. С. 13 – 20. URL: https://doi.org/10.21691/atbp.v8i1.17
9. Yavorskyi O., Tarakhtii O., Maksymov M., Kryvda V.. Model of gas turbine plant with

concentrated parameters for analysis of dynamic properties patterns. *Energy Engineering and Control Systems.* 2023. V. 9. No. 2. P. 105 – 118. URL: https://doi.org/10.23939/jeecs2023.02.105
10. Березльов В. П., Гвоздецький І. І., Капітанчук К. І. Системи автоматичного керування газотурбінних установок і компресорів К.: НАУ-ДРУК, 2010. 164 с.
11. ГОСТ 13109-97. Електрична енергія. Сумісність технічних засобів електромагнітна. Норми якості електричної енергії в системах електропостачання загального призначення: [Введ. 01.01.2000]. Вид. офіц. Київ, 1998.
12. Харабет О.М. Вивчення класичної теорії автоматичного управління за допомогою сучасного персонального комп'ютера. Одеса: Бахва, 2014. 188 с.

IMPROVEMENT OF THE AUTOMATIC CONTROL SYSTEM OF A COGENERATION POWER PLANT BASED ON A GAS TURBINE ENGINE

O. E. Mishkoy, O. S. Tarakhtij

National Odesa Polytechnic University
1, Shevchenka Ave., Odessa, 65044, Ukraine
Email: mishkoy2002@gmail.com

The article deals with the issue of improving the quality of power control of a gas turbine unit (GTU), which is a thermal engine of a cogeneration power plant. The relevance of the topic is due to the difficult economic situation in the energy sector of Ukraine, high prices for imported energy resources and the need for rational use of local and alternative fuels. The paper emphasizes the importance of improving the energy efficiency of the GTU, in particular through control automation, which allows for stable operation of the unit under various external conditions and fuel types. The main goal of the work was to develop the structure of the automatic power control system for GTUs, which takes into account external disturbances, such as changes in electrical load, fuel combustion heat, and ambient air temperature. It is proposed to use an additional signal based on the derivative of the temperature of the working fluid in front of the turbine, which contributes to a significant improvement in transient performance. The article presents a structural diagram of automatic control, simulates transient processes of turbine rotor speed control as the main regulated parameter, and performs a comparative analysis of the quality of control when an additional signal is introduced based on the derivative of the temperature of gases at the turbine inlet. The modeling results showed that the use of an additional derivative signal increases the stability of the system, reduces the dynamic deviation of parameters by up to 46.8% and improves the damping of oscillations by 16%. In addition, the introduction of automated control with a derivative signal reduces the deviation of the gas temperature at the outlet of the combustion chamber, which reduces the thermal stress on the elements of the flow part of the GTU.

Keywords: gas turbine units, cogeneration, automatic control, transient modes, modeling, external disturbances.

DOI 10.15276/imms.v14.no4.366
УДК 004.421.4:004.02

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 366-379

РОЗРОБКА ЗАСТОСУНКУ ДЛЯ БІОМЕТРИЧНОЇ ВЕРИФІКАЦІЇ ПІДПИСІВ

Р. І. Назаренко[1], О. А. Стопакевич[1], А. О. Стопакевич[2]

[1]Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
[2]Державний університет інтелектуальних технологій
1, Кузнечна вул., Одеса, 65000, Україна
Email: stopakevich@gmail.com

Верифікація підпису - це процес, під час якого заданий підпис порівнюється з деякими відомими зразками підписів і вирішується, чи належить підпис, що перевіряється, тому самому автору чи ні. Ця перевірка здійснюється за допомогою відповідного програмного забезпечення, створеного за результатами проведених досліджень. Таким чином, метою роботи є розробка алгоритму для ефективного розв'язку задачі верифікації підписів, створення програмного застосунку на основі розробленого алгоритму, перевірка працездатності системи шляхом дослідження функціонування розробленого застосунку з використанням достатньої бази даних автентичних та підроблених підписів різних підписантів. Проведено аналіз стану проблеми, який дозволив знайти базу з 500 автентичних та підроблених підписів, визначити 15 основних ознак підробки підпису як такого, 10 математичних методів порівняння підписів, основні характеристики та структуру алгоритму функціонування застосунку, його архітектуру. Експериментальні дослідження, проведені для перевірки знайдених ознак підробки підпису як такого за числовим значенням критерію стабільності ознаки, дозволили виявити ті ознаки, які доцільно використати в застосунку для перевірки. Також, експериментальні дослідження на базі підписів, проведені для перевірки знайдених математичних методи порівняння відомого автентичного підпису та підпису, який перевіряється, дозволили виявити методи порівняння, які доцільно використати в застосунку. При функціонуванні застосунку цифрові зображення підписів, отримані за допомогою цифрової камери або сканера, обробляються й за допомогою методів розпізнавання образів, знайдених ознак та методів порівняння, здійснюється перевірка. На використаній базі результати перевірки виявилися успішними. Розроблений застосунок може бути удосконалений та послугувати основою економної системи верифікацій підписів.
Ключові слова: біометрична верифікація підписів, кібербезпека, комп'ютерний застосунок, розпізнавання образів.

Вступ. Підпис – це характерно стилізований, написаний від руки варіант чийогось імені або іншого ідентифікаційного слова чи символу, який може бути використаний для підтвердження особи відповідної особи. Підпис людини є характерно унікальний і візуально розрізнюваний. Це зробило підпис давнім розпізнавальним знаком для ідентифікації особи. Навіть сьогодні власноручний підпис є найпоширенішим засобом засвідчення автентичності будь-яких офіційних або фінансових документів. Підпис відображує певні психологічні особливості людини. Сучасні вчені називають сім основних характеристик почерку, за якими можна створити портрет людини: розмір літер, їхній нахил і форма, напрямок почерку, інтенсивність натиску, характер написання слів і загальна оцінка почерку. Верифікація підпису - це процес, під час якого заданий підпис порівнюється з деякими відомими зразками підписів і вирішується, чи належить підпис, що перевіряється, тому самому автору чи ні. Ця перевірка здійснюється за допомогою певних характерних досліджень або аналізу. При автоматичній перевірці підписів перевірка підписів здійснюється машиною з використанням деяких методів розпізнавання образів. Під час перевірки за зображенням підписів зображення підписів отримують за допомогою цифрової камери або сканера. Ці цифрові зображення підписів

обробляються далі й за допомогою методів розпізнавання образів здійснюється перевірка. Мета роботи – провести аналіз стану наукової проблеми, розробити алгоритми для ефективного розв'язку задачі верифікації підписів, упевнитись в працездатності алгоритмів шляхом розробки прототипу застосунку.

Аналіз стану наукової проблеми. Для практичного використання біологічні ознаки при ідентифікації особи повинні відповідати такими основним вимогам [1-10]: універсальність, тобто ознака має розпізнаватися у кожному підпису і не має втрачається через нещасний випадок або хворобу; відмінність, тобто або ознаки у підписів двох різних людей повинні бути достатньо відмінними, щоб надійно відрізнити одну підпис від іншої; інваріантність, тобто ознака має бути стабільною протягом тривалого періоду часу; простота вимірювання, бажано без застосування спеціального обладнання; та необхідно враховувати наступні фактори [3]: швидкість розпізнавання; зручність для користувачів; захищеність, тобто захист від подачі фальсифікованої інформації. Підпис - це стилізований, написаний від руки варіант чийогось імені або іншого ідентифікаційного слова чи символу, який може бути використаний для підтвердження особи. Він є унікальним та візуально розрізнюваним. Навіть сьогодні власноручний підпис є найпоширенішим засобом засвідчення автентичності будь-яких офіційних або фінансових документів [4]. Основними причинами популярності підпису є простота виготовлення; перевірка власноручного підпису здійснюється візуально без особливих труднощів; підпис може бути переоформлений, тобто за потреби (у випадку компрометації) можна змінити свій підпис до певної міри, що неможливо з іншими біометричними даними [10]. Сучасні вчені називають сім основних характеристик почерку: розмір літер, їхній нахил і форма, напрямок почерку, інтенсивність натиску, характер написання слів і загальна оцінка почерку [5, 7]. Графологи вважають, що людський мозок підсвідомо "водить" рукою того, хто пише. Цим пояснюється і те, що під час дорослішання і зміни характеру почерк людини змінюється. Однак зв'язок не є достатньо надійним для того, щоб робити впевнені висновки. Мета-аналіз понад 200 досліджень показав, що графологія виявилася нездатною визначити наявність будь-якої риси особистості, що виявляється за будь-якою методикою тестування. Також графологам не вдалося правдиво оцінити й трудові здібності людини. [8, 9]. Перелік основних характеристик підписів з позиції графології показано на рис. 1.

Рис.1. Класифікація ознак підписів [11,12]

Рис.2. Оригінальні підписи [13, 14]

Суттєво більшого прогресу досягнуто у вивченні почерку під час досудового розслідування кримінальних справ, яке має в основному криміналістичні завдання [5-7]. Почеркознавча експертиза визначає, ким виконано текст досліджуваного документа,

одна чи різні особи писали в різних документах, чи є справжнім підпис тощо. Процедури кримінальної почеркознавчої експертизи підписів показують, що лише за формальними ознаками, показаними на рис. ., визначати правдивість підпису не коректне.

Це пояснюється такими основними причинами: залежно від психофізичного стану людини та її соціокультурного оточення, в різний час підписи однієї й тієї ж особи можуть суттєво відрізнятися [15]; дві справжні підписи однієї особи ніколи не можуть бути геометрично ідентичними; не всі підписи можуть бути класифіковані за приведеними на рис.1. ознаками, а підписи можуть бути складними для класифікації, наприклад ті, що показані на рис. 2. Інші підписи можуть бути занадто простими. В криміналістиці орієнтуються на те, що підпис є психологічним відображенням стану людини. Тому для людини її підпис є природним, а для інших людей виконання чужого підпису вимагає додаткової роботи, сліди якої й шукаються. Признаками підробки можуть бути: порушення координації рухів, немотивовані зупинки, уповільнений темп письма, вдавлені штрихи, неприродне натиснення, надмірна механічно сформована рівномірність елементів підпису тощо. Зазвичай, дослідники розглядають класифікацію підробок підпису, приведену в табл. 1.

Таблиця 1.

Класифікація підробок підпису

№	Деталізований опис підробки підпису [16]	Звичайна назва
1	Підпис є справжнім підписом іншої особи	Випадкова
2	Підробляється, знаючи лише ім'я справжнього підписанта	Проста чи випадкова
3	Підпис візуально імітується	Проста
4	Виготовлена без знання правопису підпису	Проста
5	Виготовлена без належного тренування	Проста
6	Виготовлена професійним фальсифікатором	Кваліфікована

Верифікація підпису - це процес, під час якого заданий підпис порівнюється з деякими відомими зразками підписів і вирішується, чи належить підпис, що перевіряється, тому самому автору чи ні. При автоматичній перевірці, перевірка здійснюється машиною з використанням деяких методів розпізнавання образів. Дослідження в галузі автоматичної верифікації підписів почалися з середини 60-х років 20 ст. [4]. Спочатку верифікація підписів ґрунтувалася на статичній природі підпису та використовувалися геометричні параметри, пов'язані з формою підпису, але пізніше також почали розглядати динамічні ознаки підписів, такі як початкова точка підпису, напрямок штрихів, кількість штрихів, швидкість і прискорення пера тощо. Таким чином, сформувалися два підходи [17,18]. В першому зображення підписів отримуються за допомогою цифрової камери або сканера, обробляються далі й за допомогою розпізнавання образів здійснюється перевірка. В другому використовують спеціальні пристрої, наприклад: оцифровуючий планшет, електронне перо, персональний цифровий асистент, спеціальний стілус із встановленою на ньому камерою, які генерують сигнали відповідно до динамічного відтворення підпису. Отримані параметри не є візуальними і фальсифікатор ніколи не зможе отримати до них. Тому при перевірці з додатковим пристроєм може бути потенційно досягнуто вищий рівень розпізнавання.

Недоліком такої перевірки є: необхідність залучення людини, чий підпис перевіряється; процедура не є автоматичною, оскільки від людини вимагається зробити підпис не на папері, а на спеціальному пристрої; коректна процедура вимагає не тільки зняття підпису, а й перевірку попередніх підписів зі знятим, щоб виключити фактор впливу психологічного дискомфорту й незвичності на процедуру зняття. Метод є доцільним лише для перевірки дуже важливих підписів. В статті розглянуто задачу перевірки підпису лише за зображенням.

Процедура верифікації підпису узагальнено має наступні етапи: збір даних – за допомогою цифрової камери або сканера фіксується зображення підпису (шаблон); проводиться попередня обробка: фільтрація, проріджування, бінаризація, обрізання, зміна розміру, скелетування, корекція перекосу, корекція нахилу; виділення ознак – виконується для зменшення обсягу даних, присутніх у шаблонах, шляхом вимірювання їх [19]; вибір надійних ознак шляхом відбору найкращих ознак і видалення нерелевантних ознак з повного набору ознак [20, 21]; класифікація підписів на автентичний і підроблений; ефективність оцінюється за частотою помилкових відмов (FRR) та частотою помилкових підтверджень (FAR) [22].

Перевірка підпису за зображеннями ведеться за такими підходами: зіставлення з шаблонами; статистичний; структурний або синтаксичний; на основі спектрального аналізу; на основі нейронних мереж [22–26]. У зіставленні створюється шаблон на основі наявних навчальних зображень. Цей шаблон порівнюється з новим тестовим зразком. Якщо тестовий зразок має варіації або спотворення, швидкість розпізнавання в цьому підході знижується. Це робить зіставлення з шаблоном підходом до розпізнавання образів. Тому, якщо зображення сигнатур спотворені, неправильно орієнтовані або є великі внутрішньокласові відмінності між сигнатурами, швидкість верифікації при такому підході буде дуже низькою. Але цей підхід самий ефективний для виявлення простих підробок і не підходить для виявлення кваліфікованих підробок та є найпростішим серед усіх інших підходів. При перевірці з допомогою статистичного підходу кожен патерн (тобто зображення сигнатури) представлений d-кількістю ознак, тобто розглядається як точка в d-вимірному просторі ознак. Вибір ознак повинен бути таким, щоб вектори ознак зі схожого класу займали компактні та відокремлені (від векторів ознак інших класів) області в просторі ознак, де можна легко відокремити зразки з різних класів. Популярні статистичні підходи до розпізнавання образів – це прихована марковська модель та баєсівська модель. Структурний підхід в основному використовується з іншими методами в автономній перевірці підписів, коли зображення підпису розглядається як єдине ціле. За наявності дуже великої навчальної вибірки цей підхід дає хорошу верифікацію, але його обчислювальні витрати дуже великі. При підході з використанням спектрального аналізу підписи з урахуванням кривизни розкладаються у формат з різною роздільною здатністю. Цей метод може бути застосований до різних мов. Перевагами використання нейромережевого підходу до верифікації підписів є уніфікація вилучення ознак і класифікації та гнучкість при пошуку. Кожен підхід має свої переваги та недоліки. Підходи були подані в порядку зростання вимог до вибірки даних. Оскільки пошук в інтернеті показав, що знайти в вільному доступі великі та різноманітні бази вибірок не можливо, то доцільними до дослідження є тільки дві підходи. Ми зупинимось на статистичному, як такому, який вимагає досліджень, але для досягнення задовільного результату розпізнавання вибірка може бути малою. В роботі [27] було проаналізовано вплив роздільної здатності зображення на точність функціонування системи перевірки підпису. Експеримент показав, що для якісної роботи було достатньо роздільної здатності в 150 точок на дюйм.

Розробка комп'ютерного алгоритму для верифікації підписів. В роботі використана наступна процедура обробки.

1. Фільтрація. Скановане зображення підпису може містити шум. Шум на зображенні погіршує виділення ознак і подальші процеси розпізнавання. Тому фільтрація шуму є неминучим етапом попередньої обробки при розпізнаванні образів. Було помічено, що скановані зображення зазвичай схильні до впливу імпульсного шуму, що призводить до появи випадкових чорних та білих пікселів. Медіанний фільтр ефективно видаляє цей тип шуму, зберігаючи краї зображень.

2. Бінарізація. Спочатку кольорове RGB зображення перетворюється у відтінки сірого. Існує декілька поширених методів такого перетворення. В роботі застосовано модифікований метод врахування яскравості, який використано в відомих телевізійних

стандартах SECAM, PAL, NTSC та в пакеті MATLAB Image Processing, де $I = 0.2989 \cdot R + 0.587 \cdot G + 0.114 \cdot B$. Для перетворення відтінків сірого у бінарне зображення використовується відповідне порогове значення (значення пікселя). Якщо значення пікселя у відтінках сірого перевищує порогове значення, то новому значенню пікселя присвоюється 1 (одиниця), інакше 0 (нуль). Таким чином, нове зображення матиме лише два значення пікселів '1' (що відповідає білому кольору) та '0' (що відповідає чорному кольору).

3. Обрізання. Скановане зображення підпису містить підпис і деякі білі області, що не містять підпису. Ці надлишкові ділянки видаляються шляхом обрізання зображення до прямокутника, що обмежує частину підпису.

4. Проріджування. При проріджуванні штрихи зображення підпису стають товщиною в один піксель. Але під час проріджування може бути втрачена деяка інформація про зображення підписів, наприклад, ширина штрихів.

5. Скелетування. Зберігає зв'язність сегментів підпису, які були початково з'єднані, і видаляє з зображення вибрані пікселі переднього плану. Після скелетування зображення підпису перетворюється на комбінацію деяких тонких дуг і кривих. Одним з основних способів виконання скелетування є використання процесу морфологічного проріджування, який послідовно видаляє пікселі від межі. Цей процес триває доти, доки проріджування не стане неможливим. Приклад скелетованого зображення приведено на рис. 3.

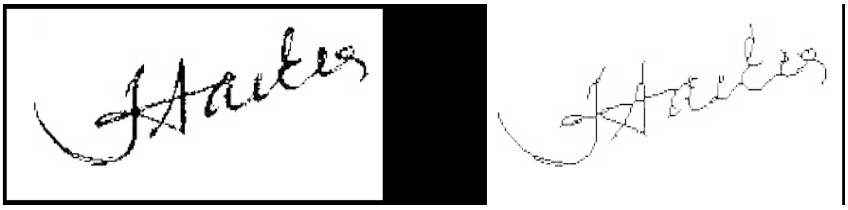


Рис.3. Приклад скелетованого зображення.

6. Корекція перекосу. Після корекції перекосу остаточне зображення робиться паралельним до горизонтальної осі. Використано метод, описаний в [28]: зображення підпису переміщають до початку координат, обчислюють мінімальне власне значення матриці, сформованої за новими координатами зображення підпису; обчислюють кут нахилу ϕ за власним вектором; після знаходження кута нахилу виконується корекція перекосу шляхом застосування перетворення обертання до кожного пікселя зображення підпису.

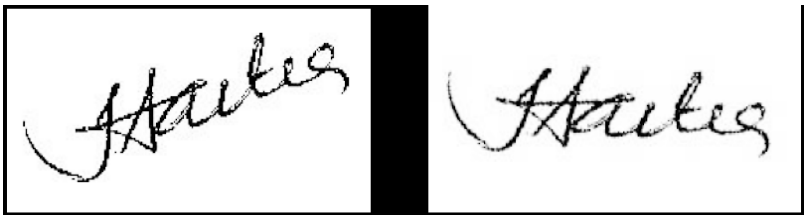


Рис.4. Приклад корекції перекосу: а-до перекосу, б-після перекосу

7. Масштабування. Довжина підпису є різною для різних підписантів. Навіть довжина підписів однієї людини також не однакова. Але коли використовується підхід перевірки підпису на основі сітки, підписи проецируються на сітку однакового розміру. Отже, всі підписи повинні бути однакового розміру. Найпростішим методом зміни розміру зображення є свого роду геометричне перетворення. Для цього є дві основні операції: просторове перетворення та інтерполяція рівня сірого. При просторовому перетворенні вибираються деякі пікселі або точки ("точки прив'язки"), положення яких на вихідному і зміненому зображенні точно відомі. На основі їхнього розташування на двох зображеннях формується рівняння просторового перетворення. Це рівняння використовується як рівняння відображення, щоб визначити положення всіх пікселів на

новому зображенні зі зміненим розміром. Інтерполяція рівня сірого використовується для присвоєння рівнів сірого новим пікселям на зміненому зображенні. Використовується метод найближчого сусіда. У цьому методі рівень сірого призначається відповідно до пікселя, який є найближчим до пікселя, що відображається [29].

Аналіз літератури показав, що в цілому для розглянутої задачі можливо визначати наступні ознаки для аналізу підпису без порівняння.

1. Нахил лінії, яка з'єднує лівий ніжній та правий верхній обрізаного зображення підпису.

2. Співвідношення сторін. Це відношення ширини підпису до висоти підпису обрізаного підпису. Видно, що співвідношення сторін підписів людини справедливо залишається постійним. Якщо висота підпису дорівнює H, а ширина підпису W, то співвідношення сторін *AR* задається формулою $AR = W / H$.

3. Чиста висота - це максимальна кількість загальних пікселів зображення (тобто чорних пікселів) серед усіх стовпців, підрахованих після обрізання зображення підпису.

4. Чиста ширина - це максимальна кількість загальних пікселів зображення (тобто чорних пікселів) серед усіх рядків, підрахованих після обрізання зображення підпису.

5. Нормалізована висота підпису - це відношення максимальної чистої висоти до максимальної чистої ширини.

6. Площа обмежувальної прямокутної рамка (реальної області підпису). Для визначення в бінарному зображення потрібно застосовувати наступний алгоритм: очищення зображення від шумових точок, що не пов'язані з основною структурою (для цього можна застосовувати Matlab функцію $im = \mathrm{bwmorph}(im_{in}, 'clean', \infty)$); визначення крайніх границь за стовпцями, а саме ліва границя $x_{\min} = \min\{i \mid \exists j : im(j,i) = 0\}$, та права границя $x_{\max} = \max\{i \mid \exists j : im(j,i) = 0\}$; пошук крайніх границь за рядками $y_{\min} = \min\{j \mid \exists i : im(j,i) = 0\}$, $y_{\max} = \max\{j \mid \exists i : im(j,i) = 0\}$; розрахунок ширини й висоти прямокутника, а саме ширина $\mathrm{w} = x_{\max} - x_{\min} + 1$, висота $\mathrm{h} = y_{\max} - y_{\min} + 1$. Таким чином, маємо координати $\mathrm{x}_{\min}, y_{\min}$ й розміри *w* і *h* прямокутника.

7. Нормалізована площа підпису (відносно обмежувальної рамки). Коли зображення підпису скелетизоване, площа підпису є мірою щільності слідів підпису. Якщо на зображенні підпису загальна кількість чорних пікселів дорівнює B, а загальна кількість пікселів у всьому зображенні дорівнює P, то нормалізована площа підпису дорівнює $NSA = B / P$. При порівнянні бінарних зображень однакових розмірів достатньо розраховувати кількість чорних пікселів.

8. Центр ваги у напряму *X*. У бінарному зображенні підпису з чорними пікселями центр ваги це координата *X* середньої точки координат усіх чорних пікселів по горизонталі

$$X = (\sum_{i=1}^{n} x_i) / N$$

де x_i - номер стовпчика чорних пікселів (Matlab функції regionprops).

9. Центр ваги у напряму *Y*. У бінарному зображенні підпису з чорними пікселями центр ваги це координата *Y* середньої точки координат усіх чорних пікселів по вертикалі

$$Y = (\sum_{i=1}^{n} y_i) / N,$$

де y_i - номер рядка чорних пікселів (Matlab функція regionprops).

10. Центр ваги X_1 лівої половини підпису у напряму *X*.

11. Центр ваги Y_1.лівої половини підпису у напряму *Y*.

12. Центр ваги X_2 правої половини підпису у напряму *X*.

13. Центр ваги Y_2. правої половини підпису у напряму *Y*.

14. Базовий зсув $Y_2 - Y_1$ (Matlab функції regionprops).

15. Нахил між центрами ваги лівої та правої половин підпису $m = (Y_2 - Y_1)/(X_2 - X_1)$.

У порядку підготовки для проведення експерименту по обґрунтуванню вибору мінімальної релевантної кількості ознак можна провести грубу класифікацію зовнішнього виду підписів на три вибірки. Дещо адаптований приклад виділення спеціальних ознак для одного типу підпису показаний на рис.5.

Рис.5. Характерні признаки підпису одного типу [30].

Таким чином, у першій вибірці підписи (300 підписів, 63% з яких є справжніми) мають нахил, діагональні та горизонтальні довгі штрихи. У другій вибірці підписи мають відірвану частину. У третьої вибірці підписи мають вертикальні довгі штрихи. Об'єми другої та третьої вибірок приблизно рівні, кількість підробок в кожній групі близько 50%. Бсього було досліджено 500 підписів. При проведенні експерименту було зроблено два істотних спрощення: варіації справжнього підпису невеликі, відмінності між справжнім та підробленим підписом візуально помітні.

Важливим є аналіз стабільності та релевантності ознак для підписів трьох різних типів. Введемо поняття індексу нестабільності, якій можна визначити за формулою

$$II = \frac{SD_G}{\mu_G} : \frac{SD_F}{\mu_F},$$

де SD_G, SD_F – стандартні відхилення ознаки автентичного та підробленого підпису, μ_G, μ_F – середні значення ознаки автентичного та підробленого підпису. Чим більше значення *II*, тим менш достовірною є ознака. Значення *II*>1 вказує на чітко нерепрезентативну ознаку, а ознаки менші за 0.5 можуть бути доцільними до розгляду.

Для кожної з трьох вибірок індекс розраховувався окремо за 15 ознаками. Автори застосували вибірку, яку отримали шляхом залучення груп людей, які робили підписи та підроблювали їх.

Результати експерименту показані на рис. 6.

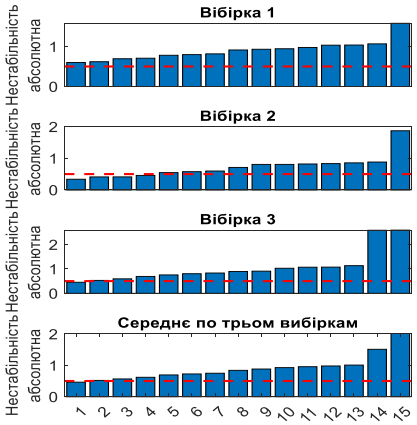


Рис. 6. Узагальнена статистика по *II* за трьома вибірками окремо та у середньому.

З рис. 6 бачимо, що у цілому майже всі ознаки мають індекс нестабільності вище за мінімально припустимий (0.5). Чим більший номер ознаки, тим більш в середньому вона не стабільна. Таким чином, виберемо ознаки з номерами 1 та 2.

Подальші експериментальні дослідження на виборці в сто підписів показали, що доцільно додати ще дві ознаки, це число Ейлера (дорівнює різниці між числом незв'язаних фрагментів підпису та числом отворів в цих фрагментах отворів в цих об'єктах, для визначення числа для бінарного зображення можна застосовувати Matlab-функцію bweuler), а також співвідношення сторін рамки підпису. Для порівняння зображення автентичних підписів необхідно спеціально підготовити, скелетузувати й привести до одного масштабу.

Далі треба визначити математичне забезпечення для порівняння різних підписів. Експериментальний аналіз ознак на вибірці в 100 підписів різних типів показав, що порівняння зображення певного реального підпису та певного підробленого підпису є задачею, яка не може бути розв'язаною з задовільною точністю.

Звичайно, можна спробувати застосувати нейромережу та машинне навчання. В прикладі [31] показано, що застосування доволі розвинутого пакета Ground DINO, який навчали визначати різні об'єкти на зображенні, не призводить до надійного результату. Час пошуку доволі тривалий (пів хвилини), а результат залежить від запиту й може бути невірним. Спеціалізоване навчання вимагає спеціальної вибірки й займе час. В той час як за допомогою визначення простих метрик задача розв'язується краще й істотно швидше.

Розглянемо задачу визначення відстаней між бінарними зображенннями *im1* та *im2*.

Індекс Жаккара $J = |im1 \cap im2| / |im1 \cup im2|$, де значення J = 1 вказує на ідентичність зображень, та його варіанти індекс Дайса $D = 2 \cdot |im1 \cap im2| / |im1| + |im2|$ та індекс Танімото

$$T = |im1 \cap im2| / \left(|im1| + |im2| - |im1 \cap im2| \right).$$

Евклідова відстань (обчислюється по пікселам) $E = \sqrt{\sum_{i,j} (im1(i,j) - im2(i,j))^2}$.

Манхеттенська відстань $M = \sum_{i,j} |im1(i,j) - im2(i,j)|$.

Хеммінгова відстань підраховує кількість відмінних пікселів між зображеннями, що відображає структурні зміни $H = \sum_{i,j} (im1(i,j) \neq im2(i,j))$.

Коефіцієнт кореляції Пірсона визначає лінійну кореляцію між зображеннями $PRS = \text{cov}(im1, im2) / \left(\sigma_{im1} \cdot \sigma_{im2} \right)$, де *PRS*=1 означає абсолютну схожість.

Логарифмічне відношення шансів $LOR = \log\left(P_1 / \left(1 - P_1\right) \right) - \log\left(P_2 / \left(1 - P_2\right) \right)$, де P_1 і P_2 - ймовірність білих пікселів для *im1* та *im2,* відображає асиметрію розподілу пікселів.

Таким чином, маємо метрики, що забезпечують різний підхід до оцінки схожості, від топологічних характеристик до геометричних і статистичних залежностей.

Виявилось що доволі позитивний результат дає порівняння нового підпису з множиною різних автентичних підписів. Для цього з автентичних підписів вибирається можливий діапазон розкиду ознак й проводиться перевірка, чи входить значення ознаки певного зображення до діапазону відхилень ознак автентичних підписів.

Для підвищення точності краще мати підписи однієї людини, що відрізняються, тобто зроблені в різний час, в різному настрої тощо. Спеціальних вимог до зображення не ставиться. Зображення має бути автоматично вичищено від фону, скелетизовано, зайві пусті частини та плями мають бути відрізані.

Порівняння зображення підпису з автентичними зображеннями підпису має проводитись в одному масштабі зі збереженням пропорцій. Збільшення кількості ознак не обов'язково призводить до гарного результату класифікації.

Насправді в більшості ситуацій це матиме негативний вплив на класифікацію. Це пов'язано з тим, що всі виділені ознаки можуть не нести суттєвої унікальності свого батьківського шаблону. Деякі з ознак несуть неоднозначну інформацію про зразок, що заплутує систему класифікатора і, як наслідок, знижує точність класифікації.

Таким чином, з усіх вилучених ознак необхідно відібрати деякі корисні ознаки для підвищення ефективності класифікації.

Якщо класифікатор наповнити нерелевантними ознаками, то можуть виникнути три проблеми: через велику кількість ознак зростають обчислювальні витрати; наявність нерелевантних ознак може призвести до помилкової класифікації і, таким чином, ефективність класифікації знижується; нерелевантні ознаки можуть спричинити надмірну підгонку.

Розроблені алгоритми, представлені на рис. 7 -11.

Рис.7. Алгоритм функціонування графічного інтерфейсу застосунку

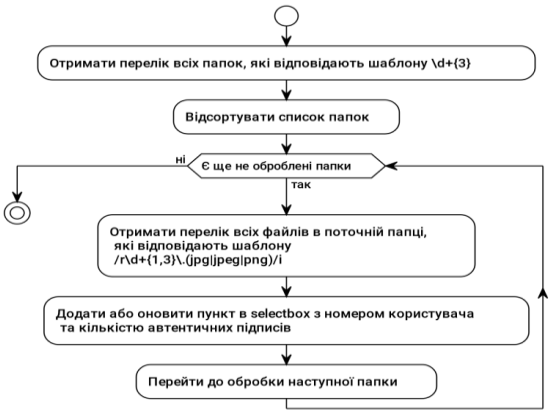


Рис. 8. Алгоритм завантаження переліку користувачів в selectbox

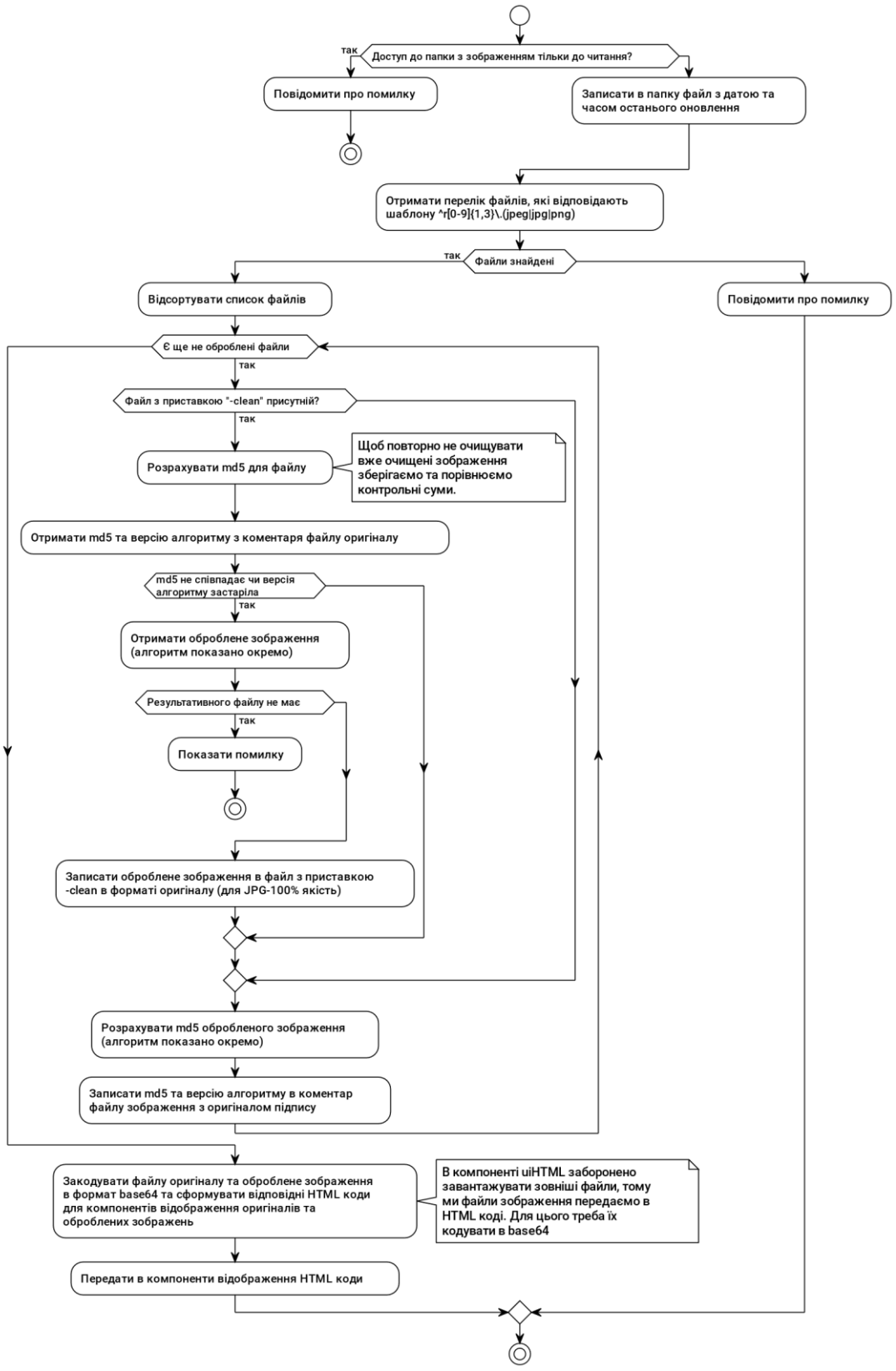


Рис.9. Алгоритм обробки зображень в папці з автентичними підписами користувача

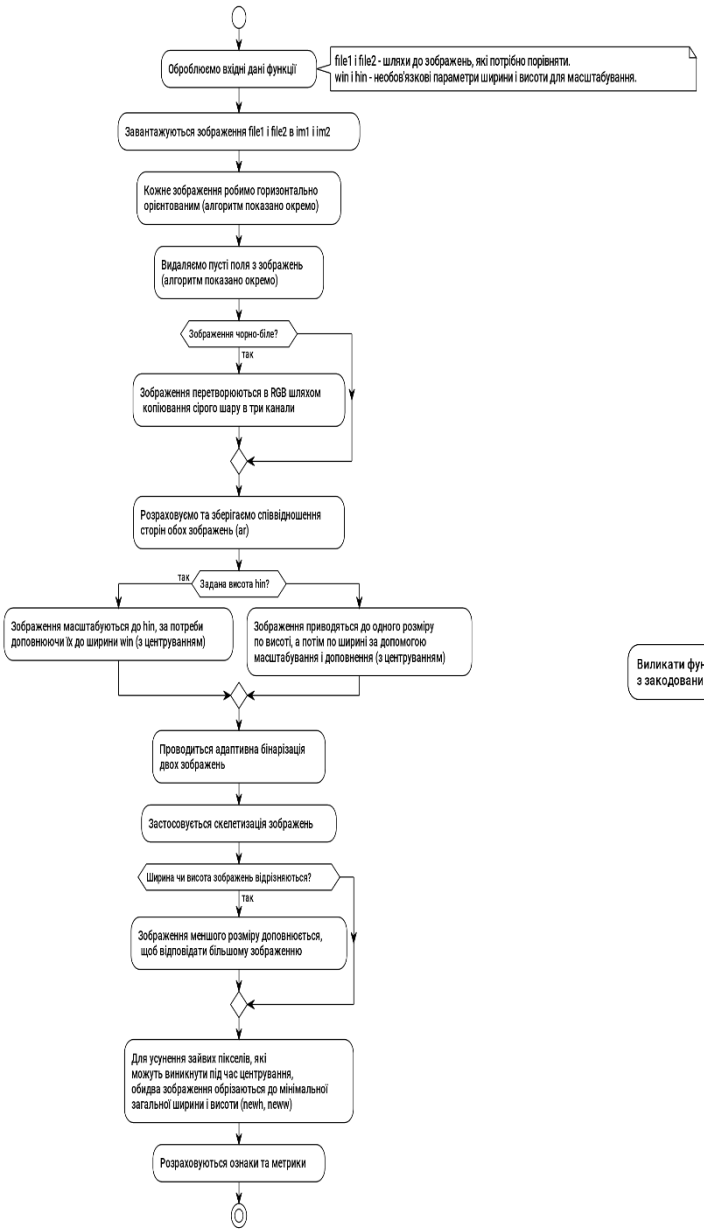


Рис.10. Алгоритм кодування зображення в форматі base64 для відображення в компоненті uiHTML

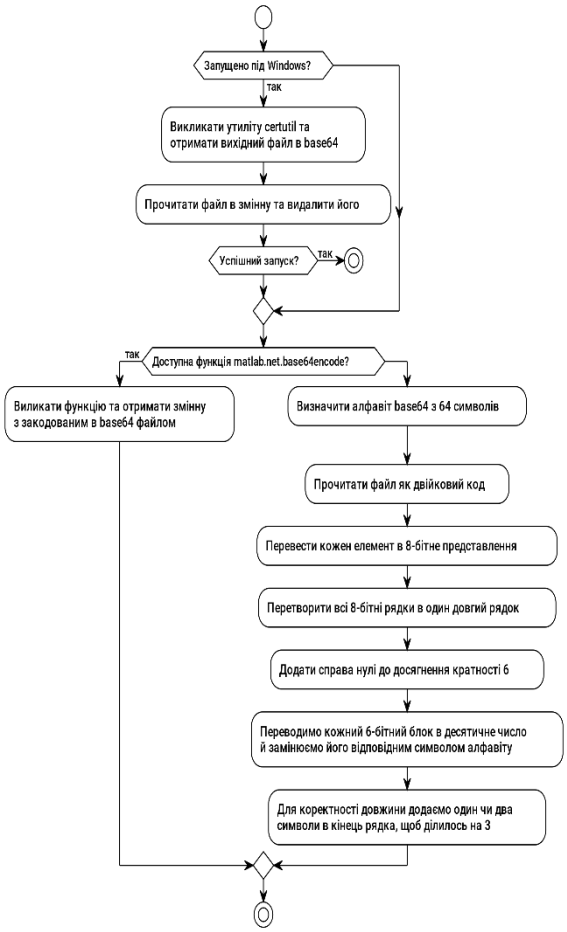


Рис.11. Алгоритм порівняння двох зображень

Розробка програмного застосунку. Розроблений макет головного вікна показаний на рис.12.

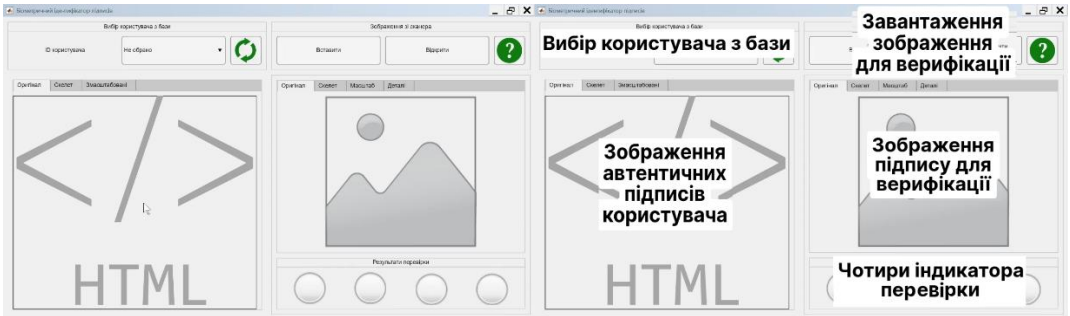


Рис.12. Макет головного вікна застосунку та поясненне його основних зон

Порядок роботи користувача застосунку з інтерфейсом наступний: з переліку клієнтів за номером обирається клієнт; візуально переглядаються реальні підписи клієнта, щоб впевнитись, що це саме той клієнт і саме той підпис (якщо підпис геть на

такий, тобто фальсифікатор не знав як виглядає реальний підпис, то перевіряти програмою не має сенсу); користувач сканує зображення для перевірки, копіює з нього область з підписом і вставляє в програму. Після перевірки її результат відображається на індикаторах. Зелений колір – індикатор пройшов перевірку, червоний – ні. Вірогідність достовірності підпису тим більша, чим більше зелених індикаторів.

Для перевірки використано набір даних під назвою "handwritten signatures - Genuine and ForgedSignature Examples" з публічного репозиторію Kaggle. Набір даних містив чотири папки зі справжніми та підробленими підписами. В результаті було сформовано 109 папок для кожного окремого користувача, в яких приблизно порівну автентичних та підроблених підписів.

Приклад результатів перевірки показаний на рис. 13.

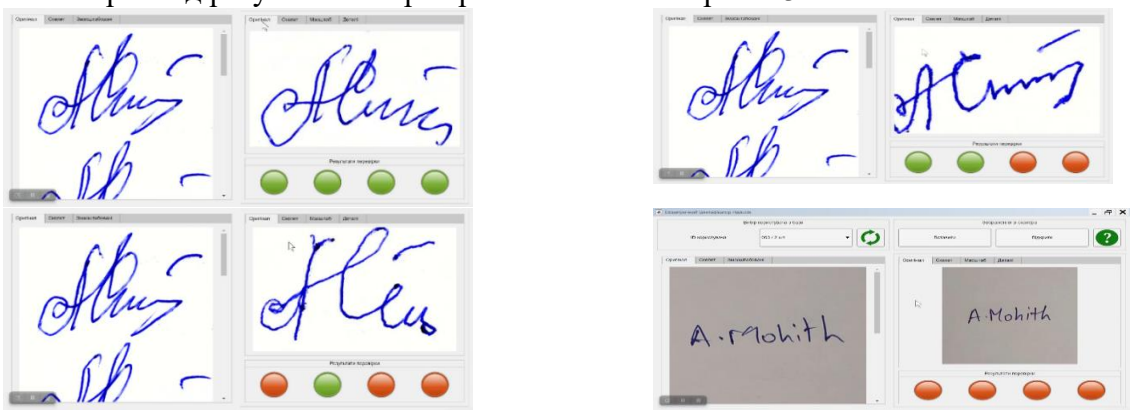


Рис.13.Приклади результатів порівняння

Отже, результат експериментальної перевірки роботи застосунку показує що робота алгоритму є задовільною.

Висновки. В роботі розв'язана задача біометричної автентифікації користувачів за підписом. Для перевірки розроблено алгоритм, який проводить аналіз скелетованих зображень, що масштабується до єдиного розміру, а також оцінюється відповідність значень признаків діапазону для автентичних підписів. Розробка застосунку здійснена на базі Matlab, оскільки ця платформа підтримує швидке прототипування й має вбудовані інструменти для обробки зображень. Програмний код застосунку реалізований у вигляді класу з графічним інтерфейсом у Matlab App Designer, що забезпечує інтуїтивно зрозумілий інтерфейс для завантаження зразків підписів, їх перегляду і порівняння. Розроблений застосунок є кросплатформовим, а інтерфейс відтворює стиль Windows для зручності користувачів. Доступна вкладка "Деталі" містить таблицю з усіма розрахованими метриками та ознаками, що підсилює можливості детального аналізу результатів. Перевірка на вибірці підписів показала високу ефективність алгоритму та підтвердила його коректність у розпізнаванні автентичних підписів та виявленні підробок.

Список літератури

1. Царьов Р.Ю., Лемеха Т.М. Біометричні технології. Одеса : ОНАЗ ім. О.С. Попова, 2016. 140 с
2. Коваль Л.Г., Злепко С.М., Новіцький Г.М., Крекотень Є.Г. Методи і технології біометричної ідентифікації за результатами літературних джерел. *Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки.* 2019. Том 30 (69). Ч. 1. № 2. С.104-112
3. Jain A.K., Ross A., Prabhakar S. An Introduction to BiometricRecognition," *IEEE Transactions on Circuits and Systems for Video Technology*. 2004. vol. 14. No.1. pp. 4-20
4. Свобода Є.Ю. Підпис як засвідчувальний знак та його значення. *Часопис Академії адвокатури України*. 2012. №16.

5. Абрамова В.М. Садченко О.О., Свобода Є.Ю. Альбом схем із судово-почеркознавчої експертизи. К.: Паливода, 2003. 120 с
6. Меленевська З.С. Свобода Є.Ю., Шаботенко А.І. Судово-почеркознавча експертиза. К.: Укр. центр духовної культури, 2007. 280 с.
7. Методика дослідження підписів. К.: ДНДЕКЦ МВС України, 2009. 21 с.
8. Dean G.A. The bottom line: effect size. Buffalo, N.Y.: Prometheus Books, 1992. p.269-341. Графологія – критичне мислення. 2023. URL: https://criticalthinkerua.wordpress.com/2023/ 07/15/graphology/
9. Pal S., Pal U., Blumenstein M. Signature-Based Biometric Authentication. In Computational Intelligence in Digital Forensics: Forensic Investigation and Applications. Springer International Publishing, 2014. P. 285-314.
10. Pinterest. URL: https://uk.pinterest.com/pin/727120302322104429/
11. Ваш підпис має значення. 2021. URL: https://vk-kp.info/novyny-ukrainy-ta-svitu/21002-vash-pidpys-maie-svoie-znachennia-i-vplyv-na-doliu-yak-zrobyty-ioho-shchaslyvym
12. Автограф як мистецтво. Незвичайні підписи знаменитостей. URL: https://www.legaltechnique.org/articles/znamenitosti/avtograf-kak-iskusstvo-neobichnie-podpisi-znamenitostej-bull-novosti-v-fotografiyah.html#google_vignette
13. Як розписуються знаменитості. URL: https://buildstuff.com.ua/yak-rozpisuyutsya-znamenitosti-pidpisi-znamenitix-lyudej-cikavi-fakti-mifi-ta-legendi/#google_vignette
14. Plamondon R. The Handwritten Signature as a Biometric Identifier:
15. Psychophysical Model and System Design. European Convention on Security and
16. Detection. 1995. P. 16-18.
17. Nguyen V., Blumenstein M., Muthukkumarasamy V., Leedham G. Off- line Signature Verification using Enhanced Modified Direction Features in Conjunction with Neural Classifiers and Support Vector Machines. *International Conference on Document Analysis and Recognition*. 2007. P. 734–738.
18. Batista L., Rivard D., Sabourin R., Granger E., Maupin P. State of the art in off-line signature verification in Pattern Recognition Technologies and Applications: Recent Advances. New York: IGI Global, 2008. P. 39-62.
19. Impedovo S., Ottaviano L., Occhinegro S. Optical character recognition – A survey. *International Journal of Pattern Recognition and Artificial Intelligence*. 1991. V. 5. No. 1/2. P. 1-24.
20. Hafemann L. G., Sabourin Ro., Oliveira L.S. Offline Handwritten Signature Verification-Literature Review. *arXiv preprint arXiv:1507.07909*. 2015. P.1-19,
21. Mauceri A.J. Feasibility studies of person identification by signature verification. Report No. SID 65 24 RADC TR 65 33. Anaheim, USA: Space and Information System Division, North American Aviation Co., 1965.
22. Wadhawan A., Kumar D. Design and Analysis of Online Punjabi Signature Verification System Using Grid Optimization. *Second International Symposium on Security in Computing and Communications, SSCC* . Delhi, India. 2014.
23. Судова експертиза: проблеми сьогодення та перспективи розвитку/ Львівський науково-дослідний інститут судових експертиз. Дрогобич : Просвіт, 2020.
24. Urmila A. Patil, J.N. Patil, N.N. Patil. A comparative study of various methods for offline signature verification. *IEEE International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad*. 2014. P. 760-764.
25. Arya M. S., Inamdar V. S. A Preliminary Study on Various Off-line Hand Written Signature Verification Approaches. *International Journal of Computer Applications.* 2010. V. 1. No. 9. P. 50-56.
26. Hou W., Ye X., Wang K. A Survey of Off-Line Signature Verification. *IEEE International Conference on Intelligent Mechatronics and Automation*. *Chengdu, China*, 2004. P. 536-541.

27. Jain A. K., Duin R. P. W., Mao J. Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2000. V. 22. No. 1. P. 4-37.
28. Theodoridis S., Koutroumbas K. Pattern Recognition. Elsevier, 2009.
29. Virajitha K., Navya B., Boggavarapu, Kumar L.N.P., Vaddi R. S. Vankayalapati H. D. Simple and Effective Techniques for Skew Correction, Slant Correction and Core-Region Detection for Cursive Word Recognition. *International Conference on Information Systems Design and Intelligent Applications. Visakhapatnam, India*. 2012. P. 353-361.
30. Gonzalez R. C.,. Woods R. E. Digital Image Processing. Prentice Hall, 2002.
31. Morales A., Morocho D., Fierrez J., Vera-Rodriguez R. Signature authentication based on human intervention: performance and complementarity with automatic systems. *IET Biometrics.* 2017. V. 6. Is. 4. P. 307–315. URL: https://doi.org/10.1049/iet-bmt.2016.0115
32. Machine learning is not all you need: a case study on signature detection. URL: https://towardsdatascience.com/machine-learning-is-not-all-you-need-a-case-study-on-signature-detection-9551f2e5d0e7

DEVELOPMENT OF AN APPLICATION FOR BIOMETRIC VERIFICATION OF SIGNATURES

R. I. Nazarenko[1], O. A. Stopakevych[1], A. O. Stopakevych[2]

[1]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
[2]State University of Intellectual Technologies
1, Kuznechna, Odesa, 65000, Ukraine».
Email: stopakevich@gmail.com

Signature verification is a process during which a given signature is compared with some known signature samples and it is decided whether the signature being verified belongs to the same author or not. This verification is carried out using appropriate software created based on the results of the research. Thus, the purpose of the work is to develop an algorithm for an effective solution to the problem of signature verification, create a software application based on the developed algorithm, and verify the system's performance by studying the functioning of the developed application using a sufficient database of authentic and forged signatures of different signatories. The article analyzes the state of the problem, which allowed finding a database of 500 authentic and forged signatures, identifying 15 main signs of signature forgery as such, 10 mathematical methods for comparing signatures, the main characteristics and structure of the application's functioning algorithm, and its architecture. Experimental studies conducted to verify the found signs of signature forgery as such by the numerical value of the stability criterion of the sign allowed to identify those signs that are appropriate to use in the application for verification. Also, experimental studies based on signatures conducted to verify the found mathematical methods for comparing a known authentic signature and the signature being verified allowed to identify comparison methods that are appropriate to use in the application. When the application is operating, digital images of signatures obtained using a digital camera or scanner are processed and verified using pattern recognition methods, found signs and comparison methods. On the database of used signatures, the verification results were successful. The developed application can be improved and serve as the basis for an economical signatures verification system.
Keywords: biometric signature verification, cybersecurity, computer application, pattern recognition, biometric signature verification, cybersecurity, computer application, pattern recognition.

DOI 10.15276/imms.v14.no4.380
УДК 621.438

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 380-390

МОДЕЛЮВАННЯ ФУНКЦІОНУВАННЯ СИСТЕМ КЕРУВАННЯ ГАЗОТУРБІННИМИ УСТАНОВКАМИ ПРИ МАНЕВРУВАННІ ЕЛЕКТРИЧНИМ НАВАНТАЖЕННЯМ

М. М. Овчинников, О. С. Тарахтій

Національний університет «Одеська політехніка»
1, Шевченка пр., м.Одеса, 65044, Україна
Email: tichyk721@gmail.com

Розподілена генерація є однією з ключових тенденцій у розвитку енергетики України. Вона передбачає створення децентралізованих джерел енергії, які розташовані ближче до кінцевих споживачів, що дозволяє зменшити втрати при передачі енергії та підвищити енергетичну незалежність. Газотурбінні установки розглядаються як перспективне рішення для забезпечення стабільності енергосистеми України в умовах значних викликів, спричинених воєнними діями та руйнуванням критичної інфраструктури. У роботі розглянуто переваги використання газотурбінних установок для розподіленої генерації, зокрема їхню адаптивність до кризових ситуацій, а також змінення динамічних властивостей газотурбінних установок у перехідних режимах в залежності від номінальної потужності установки. Досліджено роботу систем керування газотурбінними енергетичними установками в умовах змінних електричних навантажень. На основі аналізу перехідних процесів в системах керування установками при маневруванні електричним навантаженням зроблено висновок, що найкращою практикою є використання регуляторів, налаштованих на мінімально допустиме навантаження, що забезпечить гнучкість та надійність енергетичної системи.

Ключові слова: Газотурбінна установка, змінні навантаження, стабілізація енергосистеми, розподілена генерація, система керування.

Вступ. Енергетична система України зазнала безпрецедентних втрат через військові дії. Руйнування великих генеруючих об’єктів та ліній електропередач, часті атаки на критичну інфраструктуру створили загрозу для енергетичної безпеки країни. В умовах, коли централізована генерація не здатна повною мірою забезпечувати стабільність постачання електроенергії, актуальним стає розвиток розподіленої генерації. Газотурбінні установки, завдяки своїй гнучкості та ефективності, можуть відіграти важливу роль у формуванні нової, стійкої енергосистеми України. У таких умовах в нашій енергосистемі з’являється нерівномірність між виробництвом та споживанням електричної енергії. Через це виникають значні труднощі, пов’язані із дефіцитом або профіцитом електроенергії в різних регіонах нашої держави. Одним із перспективних напрямків вирішення цієї проблеми є розвиток розподіленої генерації, яка забезпечує гнучкість та надійність електропостачання. Газотурбінні енергетичні установки є достатньо маневровими і здатні працювати в умовах різких змін електричного навантаження, є можуть бути одним з основних елементів такої генерації. Отже, дослідження роботи і аналіз динамічних характеристик газотурбінних установок в умовах змінних електричних навантажень є особливо актуальною задачею в контексті відновлення та стабілізації енергосистеми України [1,2].

Руйнування великих енергетичних об’єктів унаслідок війни призвело до значного зниження доступної потужності для виробництва електроенергії. Це викликало необхідність швидкого впровадження нових рішень, здатних забезпечити енергопостачання в кризових умовах. Однак більшість традиційних підходів до генерації електроенергії потребують значного часу для відновлення. Розподілена генерація, яка базується на локальних джерелах енергії, здатна забезпечити більшу стійкість

енергосистеми. У цьому контексті газотурбінні установки мають низку переваг, але також стикаються з низкою перешкод, які потребують вирішення [1,2].

Також, слід відмітити, що у довоєнний час енергоблоки атомних електростанцій України в основному використовувалися для покриття базового навантаження у добовому графіку навантаження енергосистеми (рис.1). Ядерний реактор може працювати із заданою потужністю протягом тривалого часу тільки в тому разі, якщо на початку роботи має запас реактивності. Наразі маневрування потужністю реакторної установки здійснюється операторами в ручному режимі і тільки на вимогу диспетчерів енергосистеми. Таке виконання маневру є досить небезпечним, оскільки потребує врахування зміни багатьох нейтронно-фізичних і технологічних параметрів, а це, у свою чергу, може призвести до впливу людського фактору на безпеку АЕС [3].

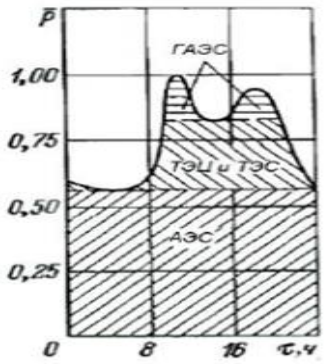


Рис. 1. Добове навантаження енергосистеми України: τ – години доби; $\overline{\boldsymbol{P}}$– споживана потужність

Проблеми використання газотурбінних установок. Газотурбінні установки здатні легко переходити на різні режими навантаження і мають значний потенціал для нівелювання вищевказаної проблеми енергетичної системи України. Їх основна перевага полягає у гнучкості та здатності до швидкого запуску, що дозволяє оперативно вводити їх в експлуатацію навіть у кризових умовах [2,3]. Це особливо важливо для регіонів, де централізована генерація недоступна через зруйновану інфраструктуру та дозволить полегшити роботу вітчизняних АЕС у маневрових режимах. Встановлення таких установок сприяє енергетичній автономії окремих територій, знижуючи залежність від централізованих мереж, які часто стають мішенню для атак. Питання використання газотурбінних установок для розподіленої генерації електричної енергії розглядаються як вітчизняними так і у зарубіжними авторами [4–13], що говорить про актуальність та великий інтерес цих типів установок.

Питання впровадження газотурбінних установок для створення розподіленої системи генерації, а також проблеми їх впровадження розглянуті у авторів [4–8]. З огляду вказаних літературних джерел можна зробити висновок, що найбільшу проблему для впровадження становить висока вартість обладнання та його інсталяції, що є значним бар’єром у контексті економічної кризи. Наступною, важливою, проблемою встановлення газотурбінних установок (ГТУ) в Україні є синхронізація роботи окремої газотурбінної установки з електромережею країни і необхідність модернізації розподільчого обладнання. Не менш суттєвим є дефіцит кваліфікованих фахівців, адже обслуговування таких систем потребує високого рівня технічної компетентності, яку через масову міграцію населення складно забезпечити [8].

Ще однією складністю є залежність від газових ресурсів. Незважаючи на потенціал альтернативного палива, більшість існуючих установок орієнтована на природний газ, що створює додаткові ризики, пов’язані з його ціною та постачанням. Крім того, пошкоджені газові магістралі та обмежена доступність до ресурсів у певних

регіонах ускладнюють їх експлуатацію. Екологічні виклики також потребують уваги, оскільки навіть сучасні ГТУ, хоча і екологічно чистіші за вугільні станції, все ще продукують викиди, які вимагають постійного моніторингу та регулювання [9–13].

Таким чином, незважаючи на значний потенціал газотурбінних установок у відновленні та модернізації енергетичної системи України, їх впровадження потребує вирішення технічних, економічних та екологічних проблем.

Особливості застосування газотурбінних установок для нівелювання нерівномірності добового навантаження. Як вже вказувалося вище, у даний час основним джерелом генерації електричної енергії в Україні – є атомні електростанції [2]. Різкі зміни навантаження для ядерних установок є, не тільки дуже несприятливими, а і небезпечними. Оскільки під час переведення енергоблоку з одного рівня потужності на другий відбувається часте переміщення регулюючої групи органів регулювання системи управління і захисту реактору, а це, у свою чергу, викликає швидке зношення і можливе руйнування цілісності оболонок ТВЕЛ реактора, а також скорочує ресурс оболонки реактора і вигоряння палива [13,14]. З метою уникнення впливу цих негативних факторів при маневруванні використовують різні методи щодо забезпечення довговічності реактору і переустановлення тепловидільних збірок для подовження ресурсу. Але все це тимчасові міри, які впливають на надійність і стійкість роботи реактору [13]. Впровадження газотурбінних установок здатне знизити навантаження на енергоблоки атомних електричних станцій під час маневрування потужністю.

Для покриття пікових навантажень застосування ГТУ є доцільним. Це пояснюється тим фактом, що питома вартість газотурбінної установки в 1,5-2 рази менша за вартість великих паротурбінних установок (ПТУ), які мають більш високий ККД. До того ж, обслуговування газотурбінних установок значно простіше, а час пуску з холодного стану за умови відповідного конструктивного виконання становить 5-15 хв. А враховуючи характерні для пікових турбін числа годин роботи на рік (500-2000 год.) застосування ГТУ виявляється рентабельним навіть незважаючи на порівняно низький ККД (0,26-0,29) і без утилізації теплоти відхідних газів [5,13].

Зазвичай сучасні ГТУ, виконані за простою схемою, виробляють як універсальні агрегати, пристосовані для різних режимів роботи. При цьому змінюється, як правило, початкова температура газів перед турбіною, і для пікового використання, зважаючи на найменше число годин роботи на рік, ГТУ експлуатується при більш високих температурах газу після камери згоряння і, отже, більш високих потужності та ККД, ніж для напівпікових і базових режимів роботи.

Робота ГТУ на часткових навантаженнях. Режими роботи газотурбінних установок (ГТУ) можуть змінюватися через ряд зовнішніх чинників. Одним із ключових аспектів є потреба адаптації потужності, яка виробляється установкою, до змін у споживанні електроенергії. Це характерно для автономних енергосистем, де ГТУ виконує функцію приводу для електричного генератора. Такий генератор працює в ізольованій мережі, забезпечуючи постачання електроенергії до споживачів. У таких умовах система повинна швидко реагувати на зміну навантаження, що створює особливі вимоги до стабільності й гнучкості роботи двигуна.

Другою важливою причиною, яка впливає на режими роботи, є зміни атмосферних умов. Температура повітря, що надходить до компресора, має вирішальне значення, оскільки вона впливає на густину повітря і, відповідно, на ефективність процесу стиску. У холодному кліматі зростає ККД, оскільки повітря стає густішим, тоді як у спекотних умовах продуктивність ГТУ може суттєво знижуватися. Крім того, атмосферний тиск і вологість також відіграють важливу роль у формуванні робочих характеристик установки.

Розглянемо принципи функціонування газотурбінного двигуна на прикладі одновальної установки, яка є найбільш розповсюдженою (рис. 2). Ця конструкція розроблена спеціально для приводу електрогенератора й забезпечує підтримання

стабільної частоти обертання ротора. Основна перевага одновальної схеми полягає у її простоті й компактності, що робить її ефективним рішенням для систем автономної генерації. Завдяки інтеграції сучасних технологій керування, такі установки здатні адаптуватися до зміни навантаження, водночас підтримуючи оптимальні параметри роботи навіть при зміні температури чи інших кліматичних факторів.

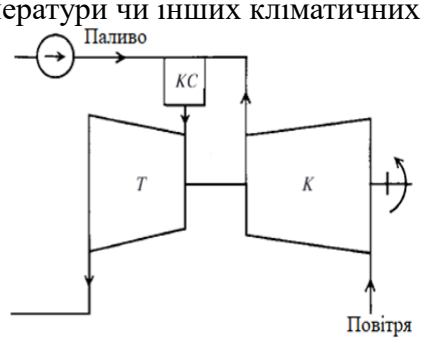


Рис. 2. Одновальна ГТУ: КС – камера згоряння; К – компресор; Т – турбіна

Для такого типу газотурбінних установок, за умови певних атмосферних умов і незмінних перетинів проточної частини компресора, камери згоряння та турбіни єдиним способом зміни навантаження є зміна витрати палива до камери згоряння.

Аналіз динамічних процесів при роботі ГТУ на часткових навантаженнях. Динамічні процеси, які виникають у газотурбінному двигуні під час зміни потужності, мають важливе значення для забезпечення його стабільної роботи. Розглянемо сценарій, у якому газотурбінна установка (ГТУ) працює в номінальному режимі, але виникає необхідність зменшити її потужність. Для цього здійснюється перекриття регулюючого клапана, що обмежує подачу палива до камери згоряння.

Зменшення витрати палива призводить до зниження температури продуктів згоряння на виході з камери. Це, у свою чергу, викликає падіння тиску газів перед турбіною, оскільки гази з меншою температурою і енергією створюють нижчий рівень тиску. Внаслідок цього зменшується і тиск повітря на виході компресора, оскільки обидва процеси тісно взаємопов’язані через конструкцію двигуна.

Однак частота обертання валу залишається постійною, що обумовлено специфікою роботи одновальних установок, розрахованих на підтримання стабільної частоти обертання генератора. Через це виникає ефект, за якого компресор змушений адаптуватися до нових умов. Згідно з витратною характеристикою компресора, що зображена на відповідних графіках (наприклад, рис. 3), зміна тиску впливає на пропускну здатність системи, і як наслідок, потік повітря через компресор дещо збільшується. Ці процеси відображають складну взаємодію між окремими компонентами ГТУ, такими як камера згоряння, турбіна й компресор. Зміна одного параметра запускає каскадний ефект, що охоплює всю систему. Для забезпечення стабільної роботи двигуна важливим є точне регулювання та оперативний контроль за його робочими параметрами, включаючи температуру, тиск і витрату повітря. Такі коригування здійснюються автоматизованими системами керування, які дозволяють підтримувати високу ефективність роботи навіть за умов зміни навантаження.

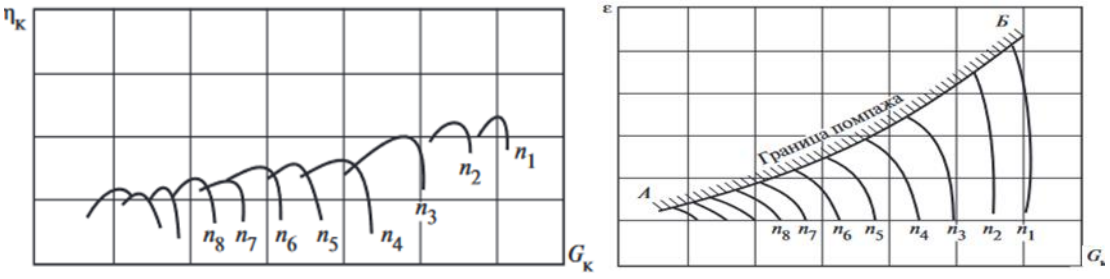


Рис. 3. Графіки зміни характеристик багатоступінчастого компресора

Таким чином, динамічні процеси в газотурбінному двигуні демонструють необхідність узгодження всіх елементів установки для досягнення оптимальних показників продуктивності й довговічності роботи в умовах змінного режиму експлуатації. Це вимагає не тільки вдосконалення конструкції окремих компонентів, але й впровадження сучасних алгоритмів керування, які забезпечують адаптацію системи до нових умов у реальному часі. Зниження витрати палива в газотурбінній установці (ГТУ) з метою зменшення її потужності супроводжується низкою взаємозалежних процесів, які безпосередньо впливають на її робочі параметри та ефективність. Основним наслідком зниження подачі палива до камери згоряння є зменшення температури продуктів згоряння, які виходять з камери та надходять до турбіни. Це зниження температури обумовлює падіння тиску газів перед газовою турбіною, що є критичним для збереження її ефективної роботи. Водночас падіння температури й тиску впливає на поведінку компресора, змінюючи його робочий режим. Згідно з характеристиками таких установок, зменшення температури газів і тиску автоматично призводить до збільшення витрати повітря. Це пояснюється тим, що компресор прагне компенсувати зниження тиску для підтримання стабільності роботи установки в умовах зменшеної подачі палива. Зменшення температури газів на виході з камери згоряння означає зниження їхньої енергії, яка використовується для приведення турбіни в рух. Відповідно, двигун працює менш ефективно, а витрата палива на одиницю виробленої енергії може зрости. Це явище стає особливо критичним у системах, які повинні працювати на змінних навантаженнях, коли часті зміни потужності створюють додаткові навантаження на всі компоненти ГТУ.
Коефіцієнт корисної дії ГТУ. У сучасних ГТУ активно використовуються системи керування, які дозволяють оптимізувати робочі параметри двигуна в реальному часі. Такі системи враховують температурні й тискові характеристики, а також потреби в потужності, забезпечуючи більш ефективну роботу установки навіть у складних умовах. Загалом, зменшення потужності ГТУ шляхом зниження витрати палива є технічно необхідним, але потребує ретельного балансу між робочими параметрами для мінімізації втрат ефективності. Цей процес ілюструє важливість комплексного підходу до управління та модернізації газотурбінних технологій.
Коефіцієнт корисної дії газотурбінної установки даного типу визначається залежністю:

$$\boldsymbol{\eta = \eta(T_c/T_a, \varepsilon, \eta_т, \eta_к)} \qquad (1)$$

де $\boldsymbol{T_c}$– початкова температура газів перед турбіною; $\boldsymbol{T_a}$– початкова температура повітря перед компресором; $\boldsymbol{\varepsilon}$ – відношення тисків в компресорі та газовій турбіні; $\boldsymbol{\eta_т}$– ККД газової турбіни; $\boldsymbol{\eta_к}$– ККД компресора. Цей вираз справедливий (з несуттєвими уточненнями) не тільки для номінального, але і для перехідного режиму. У розглянутому випадку всі параметри у виразі (1) змінюються таким чином, що це призводить до зниження ККД. Найсуттєвіший вплив на зниження ККД оказує зменшення початкової температури газів $\boldsymbol{T_c}$ і відношення тисків ε. Вплив цих параметрів на коефіцієнт корисної дії відображено на рис. 4.

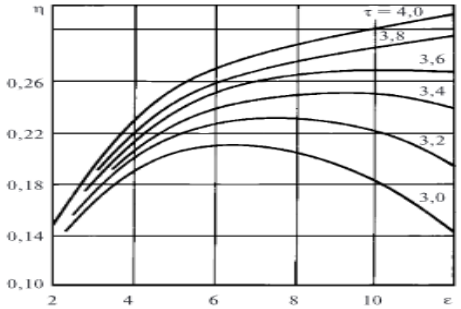


Рис. 4. Графік залежності ККД ГТУ від відношень тисків і температур

При часткових навантаження газотурбінної установки також знижуються значення коефіцієнтів корисної дії газової турбіни $\boldsymbol{\eta_т}$ і компресора $\boldsymbol{\eta_к}$, що в цілому знижує ККД всієї установки ГТУ. Окрім зміни ККД газотурбінної установки, яке відбувається під час

роботи установки у перехідних режимах, а також на часткових навантаженнях, викликає інтерес стійкість перехідних процесів регулювання при роботі ГТУ на вказаних режимах. Динаміка ГТУ при роботі на часткових навантаженнях. Основним завданням системи управління ГТУ є забезпечення стабільної роботи двигуна при заданій потужності та постійній частоті обертання ротора, незалежно від змін у зовнішніх умовах. Для цього використовуються сучасні системи автоматичного регулювання, які аналізують робочі параметри в реальному часі й вносять необхідні коригування. Це дозволяє досягати високої надійності роботи навіть у найскладніших умовах. Системи управління та захисту сучасних газотурбінних установок повинні забезпечувати виконання кількох ключових завдань. По-перше, це автоматичне і дистанційне керування основними процесами, такими як запуск, регулювання навантаження та зупинка установки. По-друге, підтримка технологічних параметрів на розрахунковому рівні. А саме, забезпечення стабільної частоти обертання турбогенератора, контроль температури газів перед і після турбіни, регулювання активного навантаження генератора та підтримання режиму роботи компресора на безпечній відстані від межі помпажу. Окрім переліченого вище, системи автоматичного управління ГТУ повинні забезпечувати виконання функції захисту в разі аварійних ситуацій. Зокрема, це запобігання перевищенню допустимої частоти обертання ротора, критичному підвищенню температури газів на вході до турбіни, зниженню тиску мастила для підшипників, згасання факела у камері згоряння та наближенню до межі помпажу компресора.

У статті [16] представлена модель динаміки газотурбінної установки і результати моделювання динамічних властивостей ГТУ при часткових і номінальних навантаженнях. Математична модель ГТУ представляє собою систему диференційних рівнянь, що описують динамічні властивості ГТУ під час дії основних збурень.

$$\begin{cases} \boldsymbol{B}\frac{\partial \boldsymbol{\Delta\omega}}{\partial t} + \boldsymbol{\Delta\omega} = \boldsymbol{b_1\Delta p_3} + \boldsymbol{b_2\Delta t_3} + \boldsymbol{b_3\Delta p_4} - \boldsymbol{b_4\Delta p_2} - \boldsymbol{b_5\Delta N}_{\text{E}}; \\ \text{T}_{\text{p}}\frac{\partial \boldsymbol{\Delta p}}{\partial t} + \boldsymbol{\Delta p} = \boldsymbol{T}_{\text{T}}\frac{\boldsymbol{d\Delta T_3}}{\boldsymbol{dt}} - \boldsymbol{k}_{\text{T}}\boldsymbol{\Delta T_3} + \boldsymbol{k_m\Delta m}_{\text{T}} + \boldsymbol{k_\omega\Delta\omega}; \\ \text{A}\frac{\partial \boldsymbol{\Delta t_3}}{\partial t} + \boldsymbol{\Delta t_3} = \text{a}_1\boldsymbol{\Delta\omega} + \text{a}_2\boldsymbol{\Delta m}_{\text{T}} + \text{a}_3\boldsymbol{\Delta t_2} + \text{a}_4\boldsymbol{\Delta t}_{\text{T}}; \\ \boldsymbol{T_2} = \boldsymbol{T_1}\cdot\boldsymbol{k_{t_2}}, \text{ де } \boldsymbol{k_{t_2}} = \boldsymbol{1} + \frac{\boldsymbol{\sigma_k}^{\boldsymbol{0.28}}\boldsymbol{-1}}{\boldsymbol{\eta_k}}. \end{cases} \quad (2)$$

Вказана система диференційних рівнянь включає три диференційних рівняння, які описують характер зміни частоти обертання ротора ГТУ, зміну тиску у газових об'ємах ГТУ, а також зміну температури газів на виході з камери згоряння і, останнє, четверте, рівняння, описує зміну температури повітря після адіабатичного стискання у компресорі. В роботі також наведені значення коефіцієнтів моделі для ГТУ широкого діапазону потужностей і графіки перехідних процесів при зміні електричного навантаження (рис. 5).

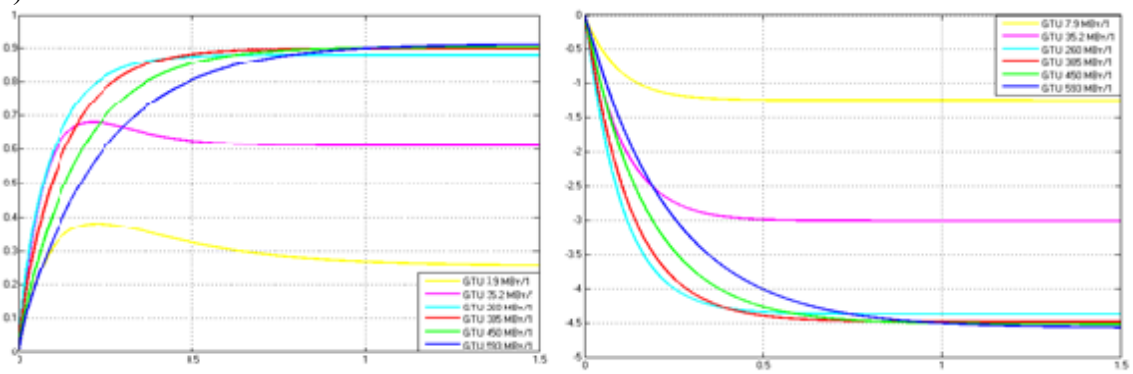


Рис. 5. Зміна частоти обертання ротора ГТУ при зміненні електричного навантаження(ΔN = ± 10%)

З представлених графіків частоти обертання ротору ГТУ видно, що найбільше

відхилення частоти спостерігається для газотурбінних установок середньої та великої потужності, причому, як при збільшенні, так і при зменшенні електричного навантаження.

Оскільки установки більшої потужності мають більшу інерційність, то із зростанням номінальної потужності газотурбінної установки динамічні властивості «погіршуються» з точки зору стабілізації частоти обертання ротора ГТУ.

Нижче будуть наведені графіки результатів моделювання перехідних процесів регулювання частоти обертання ротора ГТУ на номінальній і часткових електричних навантаженнях (75% і 50%).

Моделювання проводилося для оптимальних налаштувань регулятора, розрахованих окремо для кожного навантаження, а також за умови незмінних налаштувань регулятора і зміні електричного навантаження. У таблиці 1 наведені значення основних технологічних параметрів ГТУ, для якої проводилося моделювання.

Таблиця 1.

Номінальні значення технологічних параметрів ГТУ

$N_Г$ / Параметр	100%	75%	50%
$N_Г$, МВт	4	3	2
G_3, кг/с	10,38	7,35	6,57
G_B, кг/с	10,1	8,32	5,58
G_T, кг/с	0,206	0,166	0,117
t_3, ºС	1200	1200	1200
t_4, ºС	584	584	584

Аналіз якості керування ГТУ при зміні її електричного навантаження у широкому діапазоні. Проведено розрахунок коефіцієнтів диференціальних рівнянь моделі для навантажень 75% та 50%.

Для стабілізації частоти обертання ротора турбіни застосовується ПІ регулятор. Збуренням є зниження навантаження на 10% від номінального. При цьому відхилення частоти електричного струму в енергосистемі України не повинно перевищувати значення 0,2 $с^{-1}$[17].

Для аналізу якості перехідних процесів (ПП) були прийняті такі показниками як: перше динамічне відхилення ($\Delta\omega_1$), величина перерегулювання ($m = \Delta\omega_2/\Delta\omega_1$), ступінь загасання коливань ($\psi = (\Delta\omega_1 - \Delta\omega_3)/\Delta\omega_1$) та час регулювання ($t_p$, при якому ($\Delta\omega \le 5\%$ від $\Delta\omega_1$).

Вимоги до перехідного процесу: $\Delta\omega_1 < 0{,}2\ с^{-1}$; $m \approx 0{,}4$; $\psi \ge 0{,}85$; $t_p \le 2$ с. Оптимальні налаштуваннями регулятора, які розраховані для кожного рівня навантаження, наведені в табл.2.

Таблиця 2.

Оптимальні налаштування регулятора для номінального і часткових навантажень ГТУ

Налаштування регулятора при навантаженні	100 %	75 %	50 %
Коефіцієнт посилення регулятора $К_р$, % хро/$с^{-1}$	0,02	0,04	0,06
Час ізодрому регулятора T_i, с	2,2	1,7	1,1

Результати моделювання системи керування ГТУ при вказаних навантаженнях наведені на рис.6.

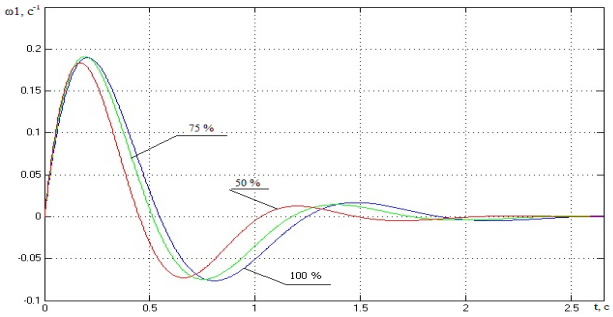


Рис. 6. Перехідні процеси при оптимальних налаштуваннях регулятора для всіх рівнів навантаження (ΔN = ± 10%)

Розраховані показники якості перехідних процесів наведені у таблиці 3.

Таблиця 3.

Показники якості перехідних процесів при зміні навантаження

Показник ПП / $N_Г$	100 %	75 %	50 %
Перше динамічне відхилення $\Delta\omega_1$, c^{-1}	0,190	0,187	0,183
Перерегулювання m	0,402	0,393	0,398
Ступінь згасання процесу ψ	0,91	0,92	0,93
Час регулювання t_P, с	1,7	1,5	1,3

З табл.3 видно, що $\Delta\omega_1$ при зниженні навантаження зберігається практично однаковою і відповідає встановленим вимогам. Результати моделювання системи керування ГТУ при зміні навантаження і з регулятором, налаштування якого розраховані для 100% навантаження, наведені на рис.7.

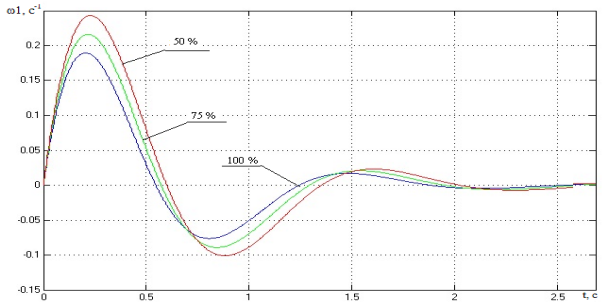


Рис. 7. Перехідні процеси регулювання частоти обертання ротора при налаштуваннях регулятора для 100% навантаження (ΔN = ± 10%)

Показники якості керування при регуляторі з налаштуванням на 75% навантаження наведені у табл. 4, з якої видно, що $\Delta\omega_1$ при 75% і 50% навантаженні перевищило допустиме відхилення у 0,2 c^{-1} на, відповідно, 14 % і 28 %, що є неприпустимим. Також спостерігається збільшення часу регулювання.

Таблиця 4.

Показники якості перехідних процесів при 100% навантаженні

Показник ПП / $N_Г$,%	100 %	75 %	50 %
Перше динамічне відхилення$\Delta\omega_1$, c^{-1}	0,190	0,216	0,243
Перерегулювання, m	0,402	0,41	0,41
Ступінь згасання процесу, ψ	0,91	0,91	0,91
Час регулювання t_P, с	1,7	1,8	1,9

Результати моделювання системи керування ГТУ при зміні навантаження з

регулятором, налаштування якого розраховані для 75% навантаження, наведені на рис.8.

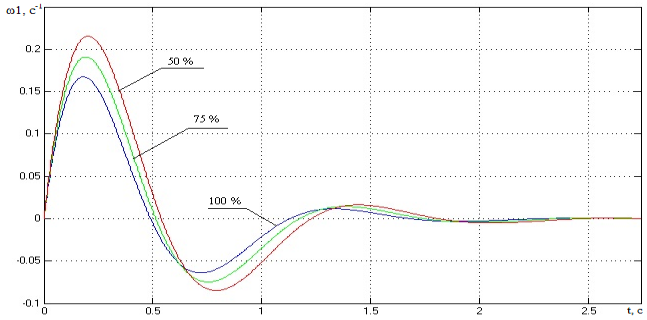


Рис. 8. Перехідні процеси регулювання частоти ротора ГТУ при оптимальних налаштуваннях регулятора для 75% навантаження

Показники перехідних процесів при оптимальних налаштуваннях регулятора для 75% навантаження наведено у таблиці 5.

Таблиця 5.

Показники перехідних процесів при 75% навантаження

Показник ПП / $N_Г$, %	100 %	75 %	50 %
Перше динамічне відхилення$\Delta\omega_1$, $с^{-1}$	0,167	0,186	0,215
Перерегулювання m	0,380	0,393	0,340
Ступінь згасання процесу ψ	0,93	0,92	0,92
Час регулювання $t_Р$, с	1,45	1,5	1,6

Встановлені налаштування регулятора не забезпечують збереження частоти обертання Δω1 у заданих межах при 50 % навантаженні. У разі зростання навантаження якість регулювання підвищується через поліпшення динамічних властивостей ГТУ.

Результати моделювання системи керування ГТУ при зміні навантаження з регулятором, налаштування якого розраховані для 50% навантаження, наведені на рис.9.

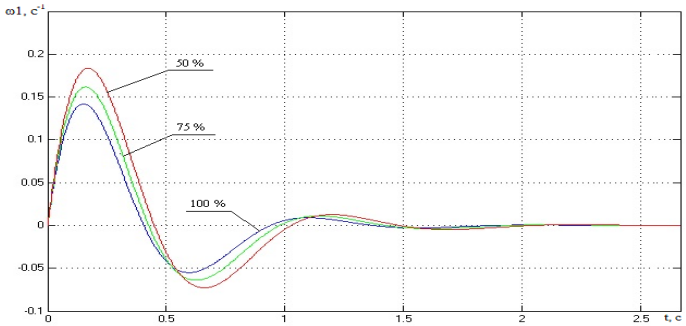


Рис. 9. Перехідні процеси регулювання частоти при оптимальних налаштуваннях регулятора для 50% навантаження

Показники якості, розраховані для вищенаведених перехідних процесів при оптимальних налаштуваннях регулятора для 50% навантаження наведено у таблиці 6.

Таблиця 6.

Показники перехідних процесів при 50% навантаження

Показник ПП / $N_Г$, %	100 %	75 %	50 %
Перше динамічне відхилення$\Delta\omega_1$, $с^{-1}$	0,142	0,162	0,183
Перерегулювання m	0,388	0,395	0,398
Ступінь згасання процесу ψ	0,94	0,93	0,93
Час регулювання $t_Р$, с	1,2	1,25	1,3

З таблиці 6 видно, що при переході газотурбінної установки на навантаження 75 % і 100 % якість регулювання зростає. При навантаженні 100% відхилення $\Delta\omega^1$ знижується на 22%, а при навантаженні 75% – на 11,5% через покращення динамічних властивостей ГТУ.

Висновок. У роботі встановлено, що при зниженні електричного навантаження газотурбінної установки у діапазоні від 100% до 50% змінюються її технологічні параметри, що негативно впливає на динамічні характеристики. Налаштування регулятора, які оптимізовані для роботи на 100% навантаження, не забезпечують належної якості регулювання частоти обертання ротора за умов 75% та 50% навантаження. Для досягнення оптимальних перехідних процесів у зазначеному діапазоні електричних навантажень бажано використовувати налаштування регулятора, які адаптовані до кожного конкретного режиму. У випадках частих змін навантаження від 100% до 50% слід застосовувати налаштування, оптимізовані для режиму 50% навантаження, щоб забезпечити стабільну роботу газотурбінної установки в усьому діапазоні.

Список літератури

1. Цьогоріч Україна значно збільшить кількість газотурбінних та біогазових установок. *Укрінформ*. 2024. URL: https://www.ukrinform.ua/rubric-economy/3685024-cogoric-ukraina-znacno-zbilsit-kilkist-gazoturbinnih-ta-biogazovih-ustanovok-ekspert.html
2. Криволап К. Українська енергосистема 2023-2024: проблеми, виклики та перспективи. URL: https://rubryka.com/blog/ukrayinska-energosystema
3. Кисельова Н.І., Погребний Я.С., Беглов К.В. Розробка автоматичної системи регулювання потужності енергоблока АЕС із ВВЕР-1000 у режимі Н. *Вчені записки ТНУ ім. Вернадського*. 2018. Т. 29(68). Ч.1. №6. С. 167–170.
4. Franco A., Casini M., Viazzo S. Gas turbines in distributed energy systems: Role, challenges, and benefits. *Renewable and Sustainable Energy Reviews*. 2020. V.119. Article 109575. P. 65. DOI: 10.1016/j.rser.2019.109575
5. Gulen S.C. Gas Turbines for Electric Power Generation. Cambridge University Press, 2022. P. 350.
6. Hale J., Kelly M. Opportunities and Challenges in the Gas Power Sector during Energy Transition. *Energy Policy*. 2021. Article 112435. P. 156-171. DOI: 10.1016/j.enpol.2020.112435
7. Gasore G., Ahlborg H., Ntagwirumugara E., Zimmerle D.. Progress for On-Grid Renewable Energy Systems: Identification of Sustainability Factors for Small-Scale Hydropower. *Energies*. 2023. V.14(4). P. 826–846. DOI: 10.3390/en14040826
8. Проблеми встановлення газотурбінних установок – синхронізація з мережею, модернізація розподільчого обладнання. URL: https://interfax.com.ua/news/economic/1033532-amp.html
9. Любименко О.М., Штепа О.А. Дослідження умов роботи та витрати палива для газотурбінної установки. *Наукові праці ДонНТУ. «Електротехніка і енергетика»*. 2020. №2(23). С.65–69. URL: https://doi.org/10.31474/2074-2630-2020-2-65-69
10. De Robbio R. Micro Gas Turbine Role in Distributed Generation with Renewable Energy Sources. *Energies*. 2023. V.16. P. 704. URL: https://doi.org/10.3390/en16020704
11. Feng D., Xiang Y., Yang J., Liao K.i, He Z. A Gas Turbine Dynamic Model Considering Power Error Correction for Distributed Cogeneration System. *IEEE Sustainable Power and Energy Conference (iSPEC)*. 2019. Beijing, China https://doi.org/10.1109/iSPEC48194.2019.8975095 .
12. World Energy Outlook 2023. URL: www.iea.org/reports/world-energy-outlook-2023
13. Gasore G., Ahlborg H., Ntagwirumugara E., Zimmerle D. Progress for On-Grid Renewable Energy Systems: Identification of Sustainability Factors for Small-Scale Hydropower in Rwanda. *Energies*. 2021. V.14(4). P. 826. URL: https://doi.org/10.3390/en14040826
14. Todortsev Yu.K., Foshch T.V., Nikolskyi M.V. Analiz metodiv upravlin-nia potuzhnistiu

enerhobloka z vodo-vodianym reak-torom pry manevruvanni. *Skhidno-Yevropeiskyi zhurnal peredovykh tekhnolohii*. 2013. №8(66). P. 3–10. URL: https://doi.org/10.15587/1729-4061.2013.19134
15. Pelykh S.N., Maksimov M.V., Gontar R.L. Principles of controlling fuel-element cladding lifetime in variable VVER-1000 loading regimes. *Atomic Energy*. 2012. No. 4(112). P. 241–249.
16. Yavorskyi O., Tarakhtii O., Maksymov M., Kryvda V. Model of gas turbine plant with concentrated parameters for analysis of dynamic properties patterns. *Energy Engineering and Control Systems*. 2023. V.9. No.2. P.105–118. URL: https://doi.org/10.23939/jeecs2023.02.105
17. ГОСТ 13109-97. Електрична енергія. Сумісність технічних засобів електромагнітна. Норми якості електричної енергії в системах електропостачання загального призначення: [Введ. 01.01.2000]. Вид. офіц. Київ, 1998.

SIMULATING THE OPERATION OF GAS TURBINE CONTROL SYSTEMS DURING ELECTRIC LOAD MANEUVERING

M. M. Ovchinnikov, O. S. Tarakhtiy

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: tichyk721@gmail.com

Distributed generation is one of the key trends in the development of Ukraine's energy sector. It involves the creation of decentralized energy sources located closer to end consumers, which allows reducing losses during energy transmission and increasing energy independence. Gas turbine plants are considered as a promising solution for ensuring the stability of the Ukrainian energy system in the face of significant challenges caused by military operations and the destruction of critical infrastructure. The paper considers the advantages of using gas turbine plants for distributed generation, in particular their adaptability to crisis situations, as well as changes in the dynamic properties of gas turbine plants in transient modes depending on the nominal power of the plant. The operation of gas turbine power plant control systems under variable electrical loads is studied. Based on the analysis of transient processes in plant control systems during electrical load maneuvering, it is concluded that the best practice is to use regulators configured for the minimum permissible load, which will ensure the flexibility and reliability of the energy system.
Keywords: Gas turbine plant, variable loads, power system stabilization, distributed generation, control system.

DOI 10.15276/imms.v14.no4.391
УДК 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 391-404

ЕКСПРЕС-АУДИТ ЯК ІНСТРУМЕНТ ОЦІНКИ ВРАЗЛИВОСТЕЙ В СИСТЕМАХ ОБРОБКИ ДАНИХ: ПІДХОДИ, МЕТОДИКИ ТА РЕКОМЕНДАЦІЇ

О. А. Сиропятов, Л. М. Тимошенко, І. В. Назарова, Н. Г. Козаченко

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044,Україна
Emails: o.a.syropiatov@op.edu.ua, l.m.timoshenko@op.edu.ua

Стаття присвячена дослідженню та розробці методики експрес-аудиту систем обробки даних на предмет захисту інформації. Основна увага приділена таким аспектам. По-перше, це порівняльний аналіз методів аудиту систем інформаційної безпеки. Проведено огляд і аналіз існуючих підходів до аудиту інформаційної безпеки з метою вибору найбільш придатних методів для використання в експрес-аудиті. Зроблено висновок, що для виконання експрес-аудиту найкраще комбінувати автоматизовані інструменти та аудит вразливостей для забезпечення ефективного виявлення загроз за мінімальний час та ресурси. По-друге, це розробка методики експрес-аудиту. Представлено підхід до швидкого отримання достовірної оцінки здатності системи обробки даних протистояти сучасним загрозам. Методика орієнтована на оперативне виявлення вразливостей та оцінку ризиків. Методика складається з послідовного ланцюжка блоків, кожен з яких містить етапи збору та обробки вхідної інформації, після чого отримані дані передаються до наступного блоку для подальшої обробки та використання. Методика дозволяє сформувати чітку структуру для оцінки поточного стану інформаційної безпеки, виявлення уразливих місць та своєчасного коригування політик безпеки з метою мінімізації потенційних ризиків. По-третє, це застосування запропонованої методики. Практично виконано експрес-аудит керуючого центру систем обробки даних підприємства, що активно проводить логістичні операції. На етапі аналізу розглянуто структуру системи, специфікацію обладнання, ліцензування та оновлення програмного забезпечення тощо. Надано практичні рекомендації щодо покращення процесів інформаційної безпеки на основі результатів експрес-аудиту. Результати дослідження показують можливість використовувати методику експрес-аудиту для швидкої оцінки стану безпеки інформаційних систем та для постійної оцінки безпеки, що дозволяє оперативно реагувати на нові загрози сьогодення.

Ключові слова: експрес-аудит, інформаційна безпека, системи обробки даних, методика аудиту, вразливості, автоматизовані інструменти, ризики

Вступ. У сучасних умовах стрімкого розвитку технологій та зростання кіберзагроз організації будь-якого масштабу опиняються перед необхідністю забезпечення надійного захисту даних. Зі збільшенням складності та витонченості кібератак традиційні підходи до інформаційної безпеки є недостатніми. У цьому контексті аудит СОД (систем обробки даних) на предмет захисту інформації є критично важливим інструментом, оскільки дозволяє не лише оцінити ефективність існуючих заходів захисту, але й виявити приховані вразливості, які можуть бути використані зловмисниками[1, 2].

В Україні через актуальні загрози сьогодення активно створюються відомчі та волонтерські системи обробки інформації, що базуються на різнотипному обладнанні, часто вживаному й отриманому через допомогу чи волонтерські ініціативи. Залучення пристроїв без повної історії експлуатації створює нові виклики для забезпечення безпеки [3-5]. За звичайних умов найкращим рішенням цієї проблеми був би традиційний аудит новостворених систем. Проте це тривалий і ресурсоємний процес, який не завжди доцільний у сучасних реаліях.

У зв’язку з цим виникає потреба у розробці доступних методик експрес-аудиту, які дозволять швидко та економічно оцінити безпеку систем обробки даних. Це

забезпечить надійний захист критичної інформації та оперативне реагування на потенційні ризики в умовах обмежених ресурсів. Отже, сьогодні важливо знайти баланс між швидкістю, точністю та економічністю аудиторських перевірок, щоб ефективно мінімізувати ризики та забезпечити належний захист даних[6, 7].

Метою дослідження є обґрунтування та розробка методики експрес-аудиту для оцінки безпеки інформаційних систем, сформованих на базі різнотипного обладнання, з урахуванням сучасних загроз та обмежених ресурсів.

Для досягнення мети необхідно вирішити наступні *завдання*.

1. Провести порівняльний аналіз основних методів аудиту СОД на предмет захисту інформації для вибору найвідповідніших в рішенні поставленої задачи.

2. Розробити методику швидкого отримання достовірної оцінки здатності системи обробки даних протистояти сучасним загрозам.

3. Виконати експрес-аудит реальної системи обробки даних за розробленою методикою.

4. Розробити рекомендації щодо використання результатів дослідження для покращення безпеки систем обробки даних.

Порівняльний аналіз основних методів аудиту СОД на предмет захисту інформації для використання в методиці експрес-аудиту. Аудит СОД на предмет захисту інформації – це процес перевірки поточного стану інформаційного захисту без впровадження змін. Це діагностичний етап, метою якого є виявлення слабких місць, оцінка ризиків та визначення відповідності стандартам. У сучасних умовах аудит СОД на предмет захисту інформації може здійснюватися різними методами, кожен з яких має свої особливості [1,4,6]. Проведемо порівняльний аналіз основних методів аудиту СОД на предмет захисту інформації для вибору найвідповідніших з подальшим використанням в методиці експрес-аудиту.

1. Ручний аудит (Manual Audit). Цей метод містить паперовий аналіз системи фахівцями вручну. Він охоплює перевірку конфігурацій, прав доступу, а також інших аспектів безпеки. Ручний аудит дозволяє детально проаналізувати складні системи та виявити вразливості, які можуть бути пропущені автоматичними інструментами. Такий підхід корисний для виявлення нестандартних або унікальних вразливостей та надає глибше розуміння специфіки системи.

2. Автоматизовані інструменти аудиту (Automated Audit Tools). Використання спеціалізованих програмних рішень для автоматизованого сканування системи на наявність відомих вразливостей. Програми, зокрема, Nessus, OpenVAS або Qualys, здатні швидко виявити слабкі місця в конфігураціях мережевих пристроїв, серверів і додатків. Цей метод підходить для регулярних перевірок у великих і складних системах.

3. Аудит конфігурацій і політик (Configuration and Policy Audit). Це оцінка конфігурацій і політик безпеки, зокрема, налаштування серверів, мережевих пристроїв і додатків. Метою є виявлення неправильних або ненадійних налаштувань, які можуть створити вразливості. Цей аудит також включає перевірку відповідності політик безпеки (наприклад, політики доступу, шифрування та використання паролів) внутрішнім і зовнішнім вимогам безпеки.

4. Аналіз журналів і подій (Log and Event Analysis). Аналіз системних журналів і записів подій для виявлення підозрілих дій, несанкціонованих спроб доступу або аномалій, які можуть свідчити про наявність вразливостей у системі. Цей метод допомагає не лише виявляти вразливості на етапі експлуатації, а також виявляти можливі атаки на ранніх стадіях.

5. Аудит вразливостей (Vulnerability Audit). Спеціалізований аудит, метою якого є виявлення відомих вразливостей у системі. Цей метод містить використання як автоматичних інструментів, так і ручних перевірок для сканування системи на

наявність слабких місць, які можуть використовуватись для атак. При цьому оцінюють також актуальність оновлень і патчів безпеки.

Критерії оцінки методів аудиту. Вибір оптимальних методів аудиту визначають за трьома критеріями, оціненими за 5-бальною шкалою:

- достовірність (чи виявляє метод усі вразливості);
- швидкість(час на проведення аудиту);
- ресурсоємність(необхідні ресурси для виконання аудиту).

Оцінки методів аудиту наведено в таблиці 1.

Таблиця 1.

Оцінки методів аудиту на основі обраних критеріїв

Метод аудиту	Достовірність (1-5)	Швидкість (1-5)	Ресурсоємність (1-5)	Загальна оцінка (1-5)
Ручний аудит (Manual Audit)	5	1	3	3
Автоматизовані інструменти аудиту	4	4	4	4
Аудит конфігурацій і політик безпеки	4	2	3	3
Аналіз журналів і подій (Log Review Audit)	3	3	3	3
Аудит вразливостей (Vulnerability Assessment Audit)	5	3	4	4

З таблиці 1 слідує, що автоматизовані інструменти аудиту та аудит вразливостей мають найвищу оцінку, що робить їх найбільш ефективними для виявлення загроз. Ручний аудит забезпечує високу достовірність, але є ресурсоємним і повільним. Отже, для виконання експрес-аудиту найкраще комбінувати автоматизовані інструменти та аудит вразливостей для забезпечення ефективного виявлення загроз за мінімальний час та ресурси.

Розробка методики швидкого отримання достовірної оцінки здатності системи обробки даних протистояти сучасним загрозам. Розробка методу експрес-аудиту для керуючого центру СОД підприємства, що активно проводить логістичні операції, має на меті швидке виявлення потенційних загроз та вразливостей. Ефективне і безвідмовне функціонування керуючого центру СОД критично важливе, оскільки будь-які збої, порушення в роботі інформаційних потоків або втрата даних можуть мати важкі наслідки для людського життя та фінансових ресурсів. Проведення експрес-аудиту дає змогу оперативно оцінити стан безпеки системи та вчасно вжити необхідних заходів для забезпечення безпеки даних і ресурсів підприємства.

1. Оцінка основних загроз та вразливостей інформаційної безпеки керуючого центру СОД.

Методика експрес-аудиту передбачає виявлення і оцінку основних загроз і вразливостей керуючого центру СОД[3,7,8]. Особливу увагу приділено таким загрозам, як несанкціонований доступ до даних, порушення конфіденційності, цілісності та доступності інформації. Важливою складовою є перевірка ефективності існуючих засобів захисту та виявлення слабких місць в інфраструктурі СОД, які можуть стати ціллю для атак зловмисників.

2. Основні етапи методики експрес-аудиту.

Методика експрес-аудиту складається з кількох етапів, кожен з яких сприяє ефективному виявленню загроз і вразливостей:

– ідентифікація об'єкту захисту інформації, аналіз основних активів СОД, які вимагають захисту;

– оцінка загроз і вразливостей, перевірка загроз і вразливостей СОД;

– аналіз ефективності існуючих засобів захисту, оцінка рівня захисту інформаційних потоків та політик безпеки в інфраструктурі;

– рекомендації щодо поліпшення рівня захисту СОД.

3. Методи та інструменти для проведення експрес-аудиту.

Для швидкого виявлення вразливостей у мережевій інфраструктурі, базах даних і компонентах СОД використовують спеціалізовані інструменти, зокрема:

– сканери вразливостей OpenVAS для виявлення слабких місць у системах;

– аналіз вразливостей виконують на платформі ELK Stack (Elasticsearch, Logstash, Kibana);

– інструменти тестування на проникнення Nmap для виявлення можливості несанкціонованого доступу;

– програмне забезпечення для моніторингу безпеки FlexNet, яке дозволяє вчасно відслідковувати інциденти.

4. Оцінка ефективності методики експрес-аудиту.

Ефективність методики оцінюється за кількістю виявлених загроз і вразливостей, а також за швидкістю їх виявлення та точністю. Важлива характеристика - це здатність методики оперативно реагувати на критичні ситуації і забезпечити своєчасне прийняття рішень для мінімізації ризиків. Оцінка містить також здатність методики забезпечити максимальну точність і швидкість при мінімальних витратах часу [7].

Методика експрес-аудиту керуючего центру СОД. Методика складається з послідовного ланцюжка блоків, кожен з яких містить етапи збору та обробки вхідної інформації, після чого отримані дані передаються до наступного блоку для подальшої обробки та використання.

1. Блок вхідної інформації. Цей блок збирає всі вихідні дані про керуючий центр СОД важливі для подальшого аналізу:

– структурна схема керуючого центру СОД - візуальне відображення маршрутів інформаційних потоків, зв'язок компонентів системи;

– специфікація обладнання - список всіх серверів, маршрутизаторів, мережевих сховищ та інших ключових компонентів, включаючи дані про виробника, версії, конфігурацію;

– інформація про ліцензії - дані про всі використовувані ліцензії для програмного забезпечення(ПЗ), операційних систем, мережевих пристроїв та оновлення ліцензій;

– сервісна служба оновлень - інформація про механізми оновлення ПЗ, автоматизацію процесів та терміни оновлень.

Алгоритм:

Крок 1. Використання автоматизованих систем інвентаризації обладнання Nmap.

Крок 2. Сканування ліцензій та оновлень за допомогою спеціалізованих рішень FlexNet.

2. Блок розподілу "ваги" елементів**,** де кожному елементу системи присвоюється "вага" у контексті забезпечення інформаційної безпеки.

Наприклад. Критичні сервери - висока вага, маршрутизатори - середня, периферійні пристрої - низька вага.

Алгоритм:

Крок 1. Використання критеріїв, заснованих на важливості елементів для бізнес-процесів, чутливості даних і рівні доступу.

Крок 2. Застосування шкали оцінок (від 1 до 5) для визначення ваги кожного компонента.

3. Блок аудиту елементів.

Цей блок проводить безпосередній аудит елементів системи, використовуючи комбінацію автоматизованих інструментів та методів аналізу вразливостей.

Автоматизовані інструменти: використання сканерів безпеки OpenVAS для виявлення вразливостей в обладнанні та ПЗ.

Аналіз вразливостей: використання платформи ELK Stack (Elasticsearch, Logstash, Kibana).

Алгоритм:

Крок 1. Сканування всіх елементів на вразливості з використанням автоматизованих інструментів.

Крок 2. Виявлення критичних вразливостей на основі ваги елементів.
Крок 3. Аналіз журналів і подій для виявлення слідів атак.

4. Блок збору результатів аудиту.

На цьому етапі збирають всі результати аудиту по кожному елементу та агрегують для отримання комплексної картини стану безпеки керуючого центру СОД:
– автоматичний збір звітів із систем сканування вразливостей;
– оцінка ризику для кожного елемента на основі виявлених вразливостей та їх "ваги";
– пріоритизація: вагоміші елементи з критичними вразливостями отримують максимальний пріоритет для усунення загроз.
Алгоритм:
Крок 1. Систематизація даних у звіті з указанням вразливостей, рівня їх критичності та потенційних збитків.
Крок 2. Створення списку пріоритетних завдань щодо виправлення.

5. Результатний блок - це фінальний блок, який готує комплексний висновок за результатами аудиту, виводить загальну оцінку рівня інформаційної безпеки керуючого центру СОД та дає рекомендації щодо поліпшення.

5.1. Оцінка загального рівня ризику: використання методу зваженої оцінки всіх виявлених вразливостей. Для того, щоб шкала оцінки сумарних ризиків мала сенс і була обґрунтована, можна використовувати наступні підходи [9, 10].

Обґрунтування через концентрацію ризиків. Ризики з високими значеннями можуть мати великий вплив на безпеку, а їх сумарне значення відображає загальний стан інформаційної безпеки. Використовуючи підхід з бальною шкалою, можна згрупувати ризики за рівнями та призначити чіткі порогові значення, які будуть вказувати на різницю в ступені загрози для системи.

Обґрунтування через практичну значущість. Встановлення порогових значень на основі практичної значущості ризиків у контексті їхнього впливу на діяльність організації, що дозволяє оцінити необхідність негайного вживання заходів для мінімізації збитків.

Розглянемо приклад шкали для сумарних ризиків. Низький рівень (0-10 балів) - ризики, які не мають значного впливу на систему, їх можна усунути в рамках звичайного контролю безпеки. Середній рівень (11-20 балів) - ризики, які можуть вплинути на систему в певних умовах, їх потрібно виправити, але вони не вимагають термінових заходів. Високий рівень (21-30 балів) - ризики, які вимагають швидкої реакції, через можливі серйозні збитки або порушення роботи системи, усунення цих ризиків має стати пріоритетом. Критичний рівень (31 і більше балів) - дуже високі ризики, які є безпосередньою загрозою для безпеки СОД. Ці ризики потрібно усунути негайно, оскільки їх реалізація може призвести до значних збитків або витоку даних. Ця шкала допомагає зрозуміти, які ризики вимагають негайної уваги, а які можна відкласти. Система оцінок і порогових значень допомагає зробити пріоритети в усуненні уразливостей прозорішими і структурованими.

5.2. Висновки про ефективність заходів захисту: оцінка поточної стратегії захисту інформації керуючого центру СОД та запропоновані заходи з її оптимізації.

5.3. Рекомендації щодо покращення: пропозиції по впровадженню додаткових заходів захисту (наприклад, покращення оновлень, використання захищених протоколів тощо).
Алгоритм**:**
Крок 1. Фінальна інтеграція всіх даних у комплексний звіт.
Крок 2. Порівняння з попередніми результатами аудиту (якщо доступні) для визначення динаміки поліпшень.

Узагальнений алгоритм реалізації експрес-аудиту наведено на рисунку 1.

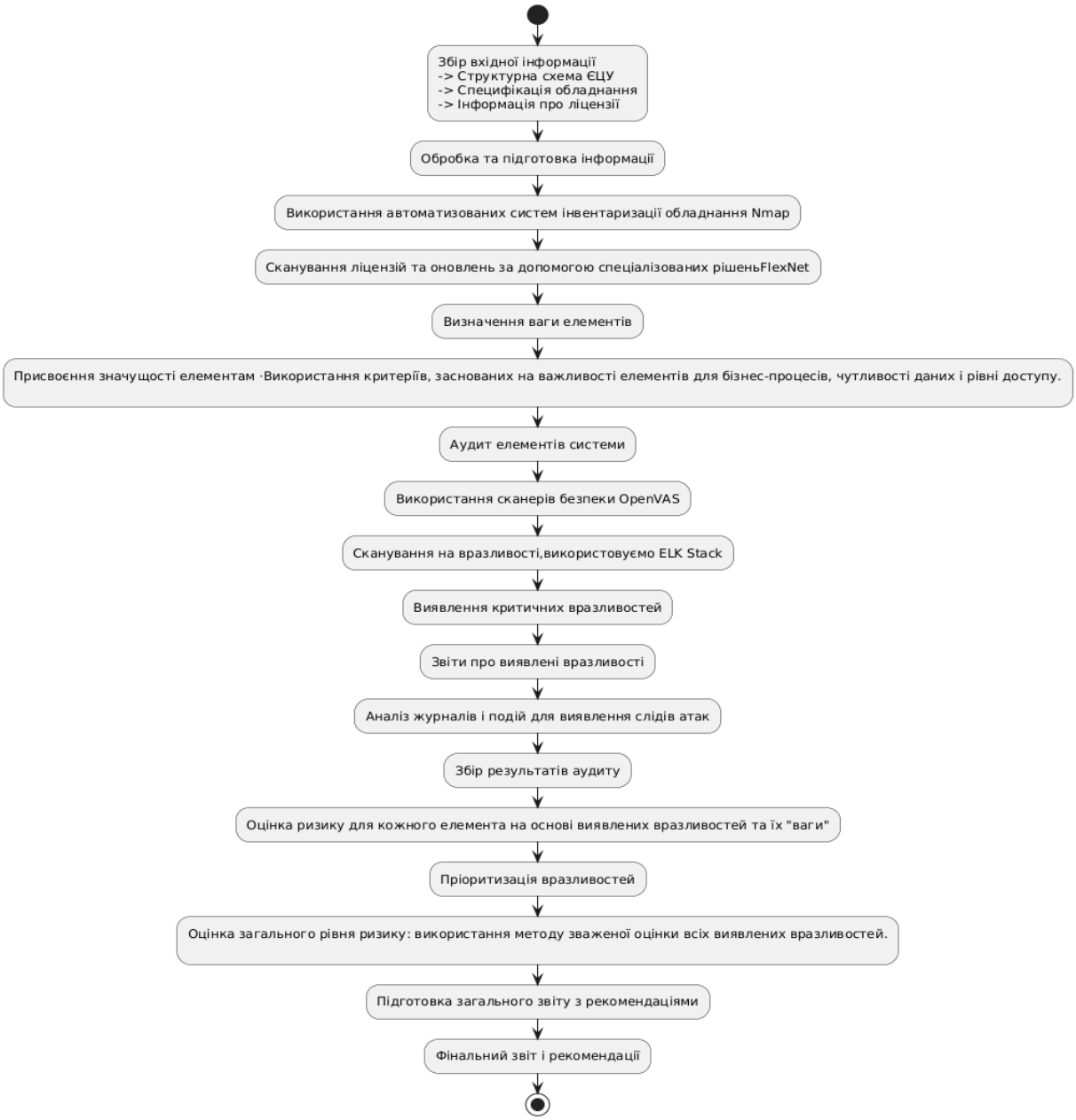


Рис. 1 Алгоритм реалізації експрес-аудиту

Отже, таким чином буде досягнута основна мета розробленої методики — визначення кроків для швидкої та достовірної оцінки здатності керуючого центру СОД протистояти сучасним загрозам. Методика дозволяє сформувати чітку структуру для оцінки поточного стану інформаційної безпеки, виявлення уразливих місць та своєчасного коригування політик безпеки з метою мінімізації потенційних ризиків.

Експрес-аудит керуючого центру СОД за розробленою методикою.

1. Блок вхідної інформації.

Збір вихідних даних про керуючого центру СОД

На цьому етапі зібрано основні дані про керуючий центр СОД, включаючи структуру, специфікації обладнання, інформацію про ліцензії та оновлення програмного забезпечення.

1.1. Структурна схема керуючого центру СОД. З метою забезпечення конфіденційності внесено деякі зміни в структурну схему та специфікації обладнання, які не впливають на хід дослідження (рисунок 2).

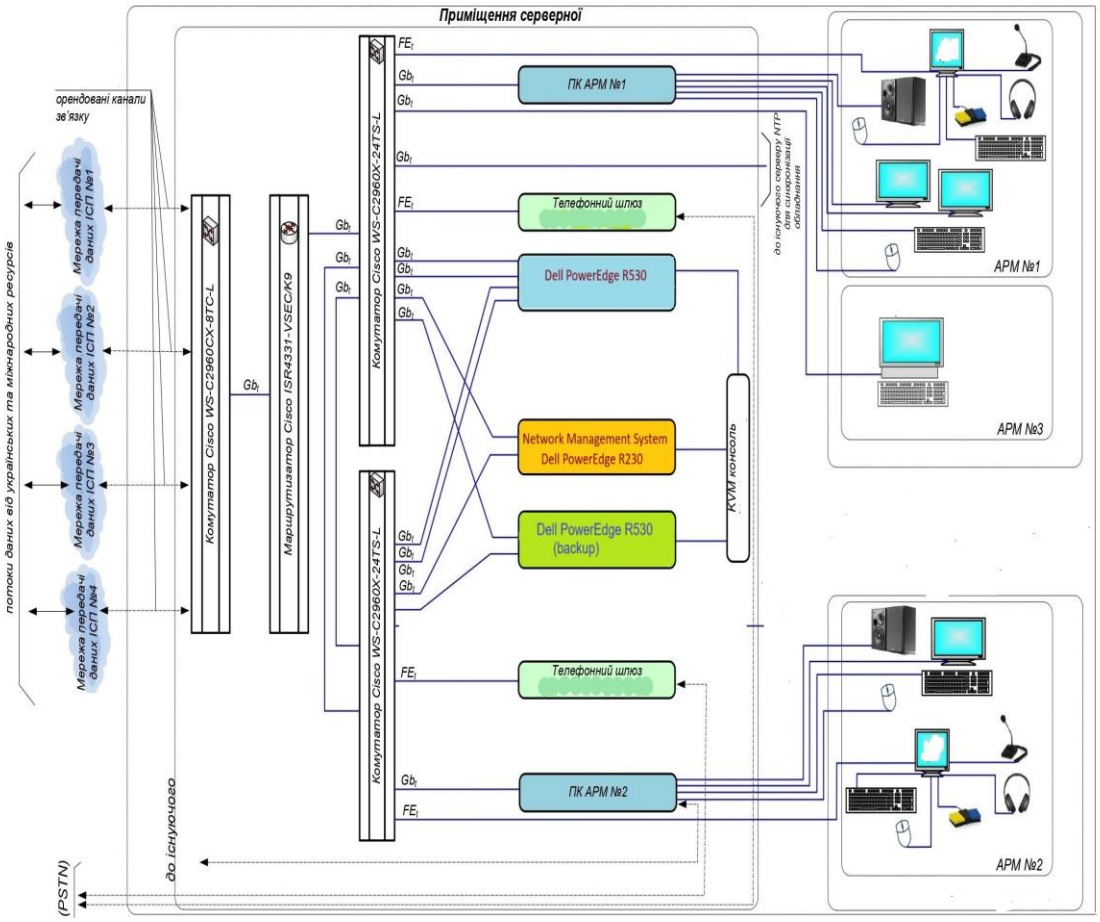


Рис.2 Структурна схема керуючого центру СОД

1.2. Специфікація обладнання:
- Сервери DELL PowerEdge R530, PowerEdge R230
- Маршрутизатор ISR4331-VSEC/K9
- Комутатори WS-C2960X-24TS-L, Cisco WS-C2960CX-8TC-L
- Сервери Supermicro SYS-5018R-M

1.3. Інформація по ліцензіях ПЗ:
- IPMI 2.0
- Dell OpenManage Essentials
- Dell OpenManage Mobile
- Dell OpenManage Power Center

1.4. Інформація по оновленням ПЗ:
- Dell SupportAssist і DELL Update Utility
- Cisco WS-C2960CX-8TC-L, WS-C2960X-24TS-L і ISR4331-VSEC/K9
- Шифрування: DES, 3DES, AES-128/256
- Аутентифікація: RSA, ECDSA
- Цілісність: MD5, SHA
- Cisco Smart Licensing та Cisco DNA Center

1.5. Заходи безпеки:
- Резервування: кластеризація серверів, резервування шлюзів.
- Шифрування даних: протоколи TLS та IPSec.
- Фізична захист: захищені приміщення з обмеженим доступом.
- Кілька провайдерів: підключення до кількох провайдерів для надійності зв'язку.

1.6. Результати застосування автоматизованих систем інвентаризації обладнання Nmap наведено в таблиці 2.

Таблиця 2.

Результати сканування з використанням автоматизованих систем інвентаризації

Найменування обладнання	Тип обладнання	Внутрішній IP-адрес*	Зовнішній IP-адрес*	Порт	Протокол	Сервіс
Комутатор Cisco WS-C2960CX-8TC-L	Комутатор	192.168.1.1	203.0.113.10	80	TCP	HTTP
Маршрутизатор Cisco ISR4331-VSEC/K9	Маршрутизатор	192.168.1.2	203.0.113.11	179	TCP	BGP
Комутатор Cisco WS-C2960X-24TS-L	Комутатор	192.168.1.3	-	80	TCP	HTTP
Комутатор Cisco WS-C2960X-24TS-L	Комутатор	192.168.1.4	-	80	TCP	HTTP
Сервер DELL PowerEdge R530	Сервер	192.168.1.5	-	443	TCP	HTTPS
Сервер DELL PowerEdge R230	Сервер	192.168.1.6	-	8080	TCP	HTTP
Телефонний шлюз PGW-1125	Шлюз	192.168.1.7	-	5060	UDP	SIP
GSM VoIP-шлюз	Шлюз	192.168.1.8	-	5060	UDP	SIP
KVM консоль ATEN CL-1000M	Консоль	192.168.1.9	-	22	TCP	SSH

*З метою забезпечення конфіденційної інформації реальні дані у стовпцях "Внутрішня IP-адреса" та "Зовнішня IP-адреса" змінені.

1.7. Результати сканування ліцензій та оновлень за допомогою FlexNet наведено в таблиці 3.

Таблиця 3.

Результати сканування ліцензій та оновлень за допомогою спеціалізованих рішень

Обладнання	ПЗ/Ліцензія	Версія	Оновлення
DELL PowerEdge R530	IPMI 2.0	Вбудована	Оновлення через iDRAC
	Dell OpenManage Essentials	Безкоштовна	Оновлення через DELL SupportAssist
	Dell OpenManage Mobile	Включена	Оновлення через DELL SupportAssist
	Dell OpenManage Power Center	Включена	Оновлення доступні на сайті DELL
DELL PowerEdge R230	IPMI 2.0	Вбудована	Оновлення через iDRAC
	Dell OpenManage Essentials	Безкоштовна	Оновлення через DELL SupportAssist
	Dell OpenManage Mobile	Включена	Оновлення через DELL SupportAssist
	Dell OpenManage Power Center	Включена	Оновлення доступні на сайті DELL
Cisco WS-C2960CX-8TC-L	Cisco IOS	15.2(4)E	Оновлення через Cisco Smart Licensing
Cisco WS-C2960X-24TS-L	Cisco IOS	15.2(4)E	Оновлення через Cisco Smart Licensing
Cisco ISR4331-VSEC/K9	Cisco IOS	16.9(4)	Оновлення через Cisco Smart Licensing

Завдяки скануванню виявлено, що деякі пристрої вимагають оновлення програмного забезпечення для підвищення рівня безпеки.

2. Блок розподілу ваги елементів

Дані щодо розподілу "ваги" елементів наведено в таблиці 4.

Таблиця 4.

Таблиця розподілу "ваги" елементів системи для забезпечення інформаційної безпеки

Елемент системи	Критичність (1-5)	Вразливість (1-5)	Наслідки порушення (1-5)	Середнє (вага)
Сервери DELL PowerEdge R530	5	4	5	4.67
Сервери DELL PowerEdge R230	4	3	4	3.67
Cisco WS-C2960CX-8TC-L	3	4	3	3.33
Cisco WS-C2960X-24TS-L	4	3	3	3.33
Cisco ISR4331-VSEC/K9	4	4	4	4.00

3. Блок аудиту елементів

Аналіз безпеки елементів керуючого центру СОД виконано за допомогою сканерів безпеки OpenVAS та аналізу журналів подій на платформі ELK.

Результат використання сканерів безпеки OpenVAS для виявлення вразливостей в обладнанні та програмному забезпеченні наведено в таблиці 5.

Таблиця 5.

Результат використання сканерів безпеки OpenVAS для виявлення вразливостей

Елемент системи	Уразливість	Оцінка уразливості (1-5)	Статус виправлення	Шифр уразливості
Cisco ISR4331-VSEC/K9	Уразливість у протоколі шифрування	3	Виправлено	CVE-2023-1234
Cisco WS-C2960CX-8TC-L	Слабкий алгоритм аутентифікації	2	Очікує виправлення	CVE-2023-2345
Cisco WS-C2960X-24TS-L	Уразливість до міжмережевих атак	4	В процесі виправлення	CVE-2023-3456
Dell PowerEdge R530	Уразливість в iDRAC	3	Виправлено	CVE-2023-4567
Dell PowerEdge R230	Недостатня захист конфіденційності	4	В процесі виправлення	CVE-2023-5678

Аналіз журналів подій на платформі ELK показує деталі активності в мережі та виявлені потенційно небезпечні події наведено в таблиці 6.

Таблиця 6.

Аналіз журналів подій на платформі ELK

Параметр	Опис	Дані за 72 години	Аномалії/Проблеми	Обладнання з аномаліями
Мережеві логи	Аналіз обсягів переданих даних та типів трафіку.	Вхідний трафік: 1,120 ГБ / Вихідний трафік:950 ГБ	Незначні сплески трафіку в нічний час.	Cisco ISR4331-VSEC/K9
Логи безпеки	Записи про спроби авторизації та доступи.	180 спроб авторизації З них 12 невдалих	Підвищена активність доступу з-за кордону.	Сервери Dell PowerEdge R530, R230
Системні логи	Повідомлення про помилки серверів та пристроїв.	28 критичних помилок на Dell R530 12 попереджень на Cisco ISR	Декілька помилок у періоди підвищеного навантаження на сервер.	Dell PowerEdge R530, Cisco ISR4331-VSEC/K9
Логи застосунків	Час відгуку застосунків, помилки баз даних.	Середній час відгуку: 450 мс Максимум: 1,000 мс	Відхилення за часом відгуку під час пікових навантажень.	Dell PowerEdge R230
Моніторинг пропускної здатності	Рівень використання каналів передачі даних.	Середнє використання: 70% Пікове використання: 92%	Часті перевантаження на каналах під час зміни потоків даних	Cisco ISR4331-VSEC/K9
Аналіз трафіку за часом доби	Піки та мінімуми у навантаженні в різний час доби.	Пік навантаження: 8-10 годин та 17-19 годин Мінімум: 02-05 годин	Навантаження перевищує оптимум у робочі години.	Cisco ISR4331-VSEC/K9

Виявлені уразливості на основі аналізу платформи ELK відображені в таблиці 7.

Таблиця 7.

Виявлені уразливості на базі аналізу платформи ELK

Ідентифікатор вразливості	Опис уразливості	Зачеплене обладнання	Критичність	Джерело даних
VULN-01	Незначні сплески трафіку в нічний час	Cisco ISR4331-VSEC/K9	Низька	Мережеві логи
VULN-02	Підвищена активність доступу з-за кордону	Dell PowerEdge R530, R230	Середня	Логи безпеки
VULN-03	Критичні помилки сервера в періоди підвищеного навантаження	Dell PowerEdge R530	Висока	Системні логи
VULN-04	Відхилення за часом відгуку при пікових навантаженнях	Dell PowerEdge R230	Середня	Логи застосунків
VULN-05	Часті перевантаження на каналах передачі даних	Cisco ISR4331-VSEC/K9	Середня	Моніторинг пропускної здатності
VULN-06	Перевищення оптимального рівня навантаження в робочі години	Cisco ISR4331-VSEC/K9	Низька	Аналіз трафіку за часом доби

Кожній уразливості присвоєно унікальний ідентифікатор (наприклад, VULN-01, VULN-02), що можна використовувати для посилань в інших розділах звіту або в рекомендаціях, що також спростить навігацію за результатами аудиту.

4. Блок збору результатів аудиту

Після виконання аудиту проведено аналіз вразливостей та рекомендацій щодо їх виправлення. Звіт з результатами наведено в таблиці 8.

Таблиця 8.

Аналіз вразливостей та рекомендації по їх виправленню

Елемент системи	**Уразливість**	**Оцінка**	**Шифр**	**Рекомендації**	**Статус виправлення**	**Примітки**	**Критичність**
Cisco ISR4331-VSEC/K9	Уразливість у протоколі шифрування	3	CVE-2023-1234	Оновити ПЗ та протоколи шифрування	Виправлено	Виправлення реалізоване у версії 15.3.2	Низька
Cisco WS-C2960CX-8TC-L	Слабкий алгоритм аутентифікації	2	CVE-2023-2345	Впровадити більш сильний алгоритм аутентифікації	Очікує виправлення	Планується виправлення в наступному оновленні	Середня
Cisco WS-C2960X-24TS-L	Уразливість до міжмережевих атак	4	CVE-2023-3456	Застосувати правила міжмережевого екрану	В процесі виправлення	Затверджено рішення, реалізація в плані	Висока
Dell PowerEdge R530	Уразливість в iDRAC	3	CVE-2023-4567	Провести аудит безпеки та оновлення	Виправлено	Успішно реалізовано, підтверджено тестами	Низька
Dell PowerEdge R230	Недостатня захист конфіденційності	4	CVE-2023-5678	Покращити конфіденційність даних	В процесі виправлення	Необхідно впровадження нових протоколів	Висока
Cisco ISR4331-VSEC/K9	Часті перевантаження на каналах передачі даних	3	VULN-05	Оптимізувати використання каналів передачі даних	Виправлено	Спостерігається покращення після змін	Середня
Dell PowerEdge R230	Відхилення за часом відгуку при пікових навантаженнях	4	VULN-04	Поліпшити продуктивність серверів	В процесі виправлення	Аналіз продуктивності триває	Середня
Dell PowerEdge R230	Уразливість у програмному забезпеченні	4	VULN-02	Оновити ПЗ до останньої версії	В процесі виправлення	Проводиться аудит версій	Висока
Cisco WS-C2960CX-8TC-L	Вразливість у механізмі авторизації	3	VULN-06	Впровадити додаткові заходи аутентифікації	Очікує виправлення	Визначено можливості для впровадження	Середня

Перелік пріоритетних завдань щодо виправлення зазначених вразливостей наведено в таблиці 9.

Таблиця 9.

Перелік пріоритетних завдань щодо виправлення

№	Елемент системи	Завдання	Причина	Термін виконання
1	Cisco WS-C2960X-24TS-L	Впровадити політики безпеки для управління доступом до обладнання. Налаштувати доступ лише для авторизованих користувачів, перевірити наявність усіх прав доступу.	Високий ризик від несанкціонованого доступу (оцінка уразливості: 4)	Протягом 3-х днів
2	Dell PowerEdge R230	Виконати шифрування даних на жорстких дисках сервера та впровадити двофакторну аутентифікацію для доступу до конфіденційних даних.	Недостатня захист конфіденційності (оцінка уразливості: 4)	Протягом 3-х днів
3	Cisco ISR4331-VSEC/K9	Оновити прошивку маршрутизатора до останньої версії, включаючи всі патчі безпеки. Реалізувати IPS для виявлення та блокування атак.	Уразливість в попередній версії (оцінка уразливості: 3)	Протягом 1 тижня
4	Cisco WS-C2960CX-8TC-L	Впровадити систему моніторингу трафіку та журналювання подій, налаштувати оповіщення про підозрілу активність.	Високий ризик від зловмисних атак (оцінка уразливості: 3)	Протягом 1 тижня
5	Dell PowerEdge R530	Провести повний аудит безпеки програмного забезпечення, впровадити регулярні оновлення та перевірки конфігурації системи iDRAC.	Уразливість в системі управління (оцінка уразливості: 3)	Протягом 1 тижня
6	Cisco ISR4331-VSEC/K9 (повторно)	Оптимізувати правила маршрутизації та налаштувати QoS для забезпечення стабільності та безпеки трафіку.	Часті перевантаження на каналах передачі даних (оцінка уразливості: 3)	Протягом 1 тижня
7	FlexNet	Провести навчання персоналу з використання ПО FlexNet для покращення управління ліцензіями та безпеки.	Недостатнє знання користувачів про функціонал ПО (оцінка уразливості: 2)	Протягом 1 місяця
8	Cisco WS-C2960CX-8TC-L	Впровадити більш сильний алгоритм аутентифікації для покращення безпеки.	Слабкий алгоритм аутентифікації (оцінка уразливості: 2)	Протягом 1 тижня
9	Dell PowerEdge R230	Поліпшити продуктивність серверів, зменшити відхилення за часом відгуку при пікових навантаженнях.	Відхилення за часом відгуку при пікових навантаженнях (оцінка уразливості: 4)	Протягом 1 тижня
10	Cisco ISR4331-VSEC/K9	Оптимізувати використання каналів передачі даних, зменшити часті перевантаження.	Часті перевантаження на каналах передачі даних (оцінка уразливості: 3)	Протягом 1 тижня

5.Результатний блок

5.1. Оцінка загального рівня ризику.

Ризики по обладнанню наведено у таблиці 10. Оцінка дорівнює 29 балів (високий рівень) — необхідно негайно усунути виявлені уразливості для уникнення важких порушень функціонування системи.

Таблиця 10.

Ризики по обладнанню

Елемент системи	Оцінка уразливості	Критичність	Ризик	Рекомендації	Статус виправлення
Cisco ISR4331-VSEC/K9	3	Середня	6	Оптимізувати правила маршрутизації та налаштувати QoS	В процесі виправлення
Cisco WS-C2960CX-8TC-L	2	Середня	4	Впровадити більш сильний алгоритм аутентифікації	Очікує виправлення
Cisco WS-C2960X-24TS-L	4	Висока	8	Впровадити політики безпеки для управління доступом	В процесі виправлення
Dell PowerEdge R530	3	Низька	3	Провести повний аудит безпеки програмного забезпечення	Виправлено
Dell PowerEdge R230	4	Висока	8	Виконати шифрування даних та впровадити двофакторну аутентифікацію	В процесі виправлення

Ризики по вразливостям наведено у таблиці 11. Оцінка дорівнює 26 балам (високий рівень) — ці вразливості можуть призвести до серйозних наслідків, тому їх усунення є пріоритетом. Ці ризики вимагають негайної уваги та дій для зменшення потенційних збитків.

Таблиця 11.

Ризики по уразливостям

Уразливість	Оцінка уразливості	Критичність	Ризик	Рекомендації	Статус виправлення
VULN-01	1	Низька	1	Моніторити сплески трафіка	Немає необхідності
VULN-02	2	Середня	4	Зменшити доступ з-за кордону	В процесі виправлення
VULN-03	4	Висока	8	Провести аудит серверних помилок	Виправлено
VULN-04	3	Середня	6	Оптимізувати продуктивність серверів	В процесі виправлення
VULN-05	3	Середня	6	Оптимізувати використання каналів передачі даних	Виправлено
VULN-06	1	Низькя	1	Аналіз трафіку за часом доби	Немає необхідності

5.2. Оцінка поточної стратегії керуючого центру СОД.

Сильні сторони - наявність базових заходів захисту, контроль доступу, резервне копіювання даних та системи моніторингу, регулярні оновлення систем, обізнаність персоналу з питань інформаційної безпеки.

Слабкі сторони - відсутність комплексної стратегії управління інформаційною безпекою, недостатня реакція на інциденти, низький рівень захисту критичних систем.

5.3. Рекомендації щодо покращення інформаційної безпеки керуючого центру СОД: впровадити автоматизовані системи для моніторингу та установки оновлень; удосконалити існуючу систему моніторингу для виявлення загроз в реальному часі; впровадити багатофакторну аутентифікацію для всіх користувачів, особливо для критичних систем; проводити регулярні внутрішні та зовнішні аудити безпеки; розробити та підтримувати актуальний план реагування на інциденти, що включає чіткі інструкції по діям при загрозах.

Рекомендації щодо використання результатів дослідження щодо покращення безпеки систем обробки даних.

1. Впровадження експрес-аудиту. Результати дослідження показують, що можливо використовувати методику експрес-аудиту для швидкої оцінки стану безпеки систем обробки даних. Це дозволить оперативно виявляти потенційні ризики та своєчасно реагувати на зміни без великих витрат.

2. Оптимізація процесу аудиту. Для підвищення ефективності можна автоматизувати деякі етапи аудиту, що зменшить час перевірки та підвищить точність результатів.

3. Регулярний моніторинг. Можливо використовувати методику експрес-аудиту для постійної оцінки безпеки, що дозволяє оперативно реагувати на нові загрози.

4. Аналіз економічної доцільності. Важливо оцінювати ефективність аудиту з точки зору витрат, щоб забезпечити баланс між високим рівнем безпеки та наявними ресурсами.

Ці рекомендації допоможуть покращити процеси безпеки та зменшити ризики без значних витрат.

Висновки. Розроблено експрес-метод виявлення та оцінки основних загроз і вразливостей у системах обробки даних. Проведено огляд та аналіз існуючих підходів до аудиту інформаційної безпеки з метою вибору найвідповіднійших для поставленого завдання. За результатами аналізу зроблено висновок, що для проведення експрес-аудиту найкраще поєднати автоматизовані інструменти та метод аудиту вразливостей, що дозволяє виявляти загрози за прийнятний час та з наявними ресурсами. На основі поєднання цих методів розроблено методику, що складається з послідовності блоків, кожен з яких містить етапи збору та обробки інформації. Ефективність методики оцінюється за кількістю виявлених загроз, швидкістю їх виявлення.

Практично використано розроблено методику для експрес-аудиту вразливостей керуючого центру СОД. Отримані результати підтвердили, що методика забезпечує реальні можливості для отримання достовірних оцінок вразливостей систем обробки даних. У статті наведено дані результатів обробки інформації кожним блоком методики під час експрес-аудиту.

Дослідження також показало, що методику можна ефективно використовувати для швидкої оцінки стану безпеки інформаційних систем і постійного моніторингу, що дозволяє оперативно реагувати на нові загрози та забезпечувати високий рівень інформаційної безпеки в реальних ситуаціях сьогодення.

Список літератури

1. Макаренко С. І. Аудит інформаційної безпеки: основні етапи, концептуальні засади, класифікація заходів. *Системи управління, зв'язку та безпеки*. 2018. № 1. С. 1–29. URL: http://sccs.intelgr.com/archive/2018-01/01-Makarenko.pdf
2. Закон України «Про основні засади забезпечення кібербезпеки України». *Відомості Верховної Ради.* 2017. № 45. URL: https://zakon.rada.gov.ua/laws/show/2163-19#Text
3. Про рішення Ради національної безпеки і оборони України від 18 березня 2022 року «Про нейтралізацію загроз інформаційній безпеці держави» № 151/2022. Верховна Рада України: офіційний вебпортал парламенту України. URL: https://zakon.rada.gov. ua/laws/show/151/2022#Text
4. Рішення «Щодо реалізації єдиної інформаційної політики в умовах воєнного стану», введено в дію Указом Президента України від 19.03.2022 року № 152/2022. Верховна Рада України: офіційний вебпортал парламенту України. URL: https://zakon.rada.gov. ua/laws/show/n0004525-22#n2
5. Борсуковський Ю. В., Борсуковська В. Ю. Прикладні аспекти захисту інформації в умовах обмеженого фінансування. *Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка»*. 2018. №1(1). С.26-34. URL: https://doi.org/10.28925/2663-4023.2018.1
6. Макаренко С.І., Смирнов Г.Є. Модель аудиту захищеності об'єкта критичної інформаційної інфраструктури тестовими інформаційно-технічними впливами. *Труди навчальних закладів зв'язку*. 2021. Т. 7. №1. С. 94–104. DOI:10.31854/1813-324X-2021-7-1-94-104

7. Огірко О.І., Крамар М.О. Аудит інформаційних систем і технологій як інструмент стратегічного управління підприємством. *Law & Sciences = Право та науки.* 2018. № 2. С.26-31.
8. Матюха М. М. Комп'ютерний аудит: опор. курс лекцій для студ. екон. спец. дистанційної форми навчання. К.: ДП «Вид. дім «Персонал», 2018. 228 с.
9. Єрмошин В.В., Невойт Я.В. Аналіз і оцінка ризиків інформаційної безпеки. Монографія. К: ДУТ, 2015. 124 с.
10. Корченко А. О. Методи ідентифікації аномальних станів для систем виявлення вторгнень. Монографія. Київ: ЦП «Компринт», 2019 . 361 с.

EXPRESS AUDIT AS A TOOL FOR ASSESSING VULNERABILITIES IN INFORMATION SYSTEMS: APPROACHES, METHODOLOGIES, AND RECOMMENDATIONS

O. A. Syropiatov, L. M. Tymoshenko, I. V. Nazarova, N. G. Kozachenko

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Emails: o.a.syropiatov@op.edu.ua, l.m.timoshenko@op.edu.ua

The article is devoted to the research and development of a methodology for express audits of data processing systems for information security. The main attention is paid to the following aspects. Comparative analysis of ISS audit methods. A review and analysis of existing approaches to information security audits is carried out in order to select the most suitable methods for use in a rapid audit. It is concluded that it is best to combine automated tools and vulnerability audits to perform a rapid audit to ensure effective threat detection with minimal time and resources. Development of a rapid audit methodology. An approach to quickly obtain a reliable assessment of the ability of a data processing system to withstand modern threats is presented. The methodology is focused on the rapid identification of vulnerabilities and risk assessment. The methodology consists of a sequential chain of blocks, each of which contains stages of collecting and processing incoming information, after which the data is transferred to the next block for further processing and use. The methodology allows forming a clear structure for assessing the current state of information security, identifying vulnerabilities, and timely adjusting security policies to minimize potential risks. Application of the methodology. An express audit of the control center of the BMS of an enterprise actively involved in logistics operations was practically performed. At the stage of analysis, the system structure, hardware specification, software licensing and updating, etc. are considered. The article provides practical recommendations for improving information security processes based on the results of the express audit. The results of the study show that it is possible to use the rapid audit methodology for a quick assessment of the security status of information systems and for continuous security assessment, which allows one to respond quickly to new threats.

Keywords: rapid audit, information security, data processing systems, audit methodology, vulnerabilities, automated tools, risks

DOI 10.15276/imms.v14.no4.405
УДК 519.6:004.93

Informatics and Mathematical Methods in Simulation
Vol.14 (2024), No. 4, pp. 405-417

МАТЕМАТИЧНІ МЕТОДИ В ОПТИМАЛЬНОМУ ВИБОРІ НАВЧАЛЬНИХ ДИСЦИПЛІН У ВИЩИХ НАВЧАЛЬНИХ ЗАКЛАДАХ

Б. І. Юхименко

Національний університет «Одеська політехніка»
1, Шевченко пр., Одеса, 65044, Україна
Email: biruteyu@gmail.com

Робота присвячена дослідженню та застосуванню математичних методів для оптимального вибору навчальних дисциплін за вибором в освітніх програмах вищих навчальних закладів. У разі загальної комп'ютеризації процеси підготовки фахівців вимагають впровадження інноваційних підходів, які забезпечують підвищення компетентності та практичності майбутніх фахівців. Розглянуто методологічні аспекти математичного програмування та експертного оцінювання, які використовуються для кількісного аналізу та вибору найбільш пріоритетних дисциплін. Особлива увага приділяється задачі оптимального вибору дисциплін за вибором із переліку запропонованих на основі критеріїв корисності та навчального навантаження. Проблема формалізується як задача лінійного програмування дискретного типу, відома як «задача про ранець». У роботі запропоновано підхід, що ґрунтується на експертних оцінках для визначення корисності дисциплін, а також побудовані математичні моделі та методи їх вирішення. Дослідження наголошує на значущості інтеграції методів експертного оцінювання у процес формування навчальних планів, що сприяє поліпшенню якості освіти та підготовки фахівців, що відповідають вимогам сучасної економіки та суспільства. Робота представляє практичну цінність для освітніх установ та розробників програм навчання, орієнтованих на потреби ринку праці.

Ключові слова: експертне оцінювання, лінійне програмування, навчальні дисципліні, освітні програми, задача про ранець

Вступ. Загальна комп’ютеризація процесів ухвалення рішень, документообігу, зберігання та передачі інформації потребує фахівців високої кваліфікації. Знання сучасної обчислювальної техніки, її математичного забезпечення, елементів штучного інтелекту є прямими вимогами до працівника будь-якого офісу. Необхідність підвищення компетентності та практичних навичок майбутніх фахівців, здатних застосовувати знання в основних і супутніх предметних галузях, є завданням вищої освіти [1-2]. Набуття знань, умінь і навичок визначається рівнем професорсько-викладацького складу, методичного та технічного забезпечення навчального процесу [3]. Комплект основних дисциплін, їх взаємозв’язок, послідовність викладення на належному методичному рівні - це важелі успішної підготовки молоді, відповідальної за добробут країни, її економіку та культуру [4].

Складання навчальних планів підготовки майбутніх фахівців здійснюється вищими органами освіти із застосуванням методик, що враховують потреби та особливості відновлення і розвитку країни. Водночас враховуються специфіка і можливості навчального закладу. До навчального плану включають дисципліни, орієнтовані на специфіку та потреби регіону.

Поглиблення знань і предметна орієнтованість майбутніх фахівців формуються завдяки дисциплінам, які включаються до навчальних планів і подаються під назвою «дисципліни за вибором». Перелік вибіркових дисциплін покликаний доповнити й

закріпити знання з урахуванням особистих побажань здобувача у відповідній предметній галузі. Вивчення «бажаних» дисциплін робить молодого фахівця впевненішим, сприяє реалізації його потенціалу і формує особистісні характеристики, необхідні для розвитку країни.

Сам процес вибору необхідної кількості дисциплін, як показує практика, є непростою процедурою. З одного боку, це організаційні вимоги, з іншого - питання, яка дисципліна буде більш корисною, якщо обрати одну з запропонованого списку. Найпростіший підхід - інтуїтивний вибір. Проте студент, який уже має базові знання, набуті впродовж певного часу, може використати їх для ухвалення обґрунтованого рішення.

Тому актуальним є створення методів та інструментів підтримки прийняття рішень [5, 6] щодо вибору дисциплін під час складання індивідуального навчального плану студента.

Мета роботи. Метою роботи є формалізація процесу вибору дисциплін за вибором для складання навчального плану підготовки майбутніх фахівців шляхом застосування математичних методів, зокрема методів оптимізації та експертного оцінювання, що дозволить забезпечити обґрунтованість і ефективність прийняття рішень.

Це сприятиме підвищенню якості освіти, врахуванню індивідуальних інтересів студентів та відповідності навчального процесу потребам регіону й сучасного ринку праці.

Для досягнення мети необхідно вирішити наступні задачі: огляд існуючих підходів до складання навчальних планів у вищих навчальних закладах з урахуванням дисциплін за вибором; розробка математичної моделі оптимального вибору дисциплін; верифікація моделі оптимального вибору дисциплін.

Постановка проблеми. Для розв’язання задачі прийняття рішень щодо вибору «дисциплін за вибором» для складання навчального плану використовують різні методи та підходи. Так в роботі [6] авторами вдосконалено евристичний багатокритеріальний метод прийняття рішень SMART і на його основі розроблено мобільну систему підтримки прийняття рішень. Застосування багатокритеріальних методів прийняття рішень (Multiple Criteria Decision-Making (MCDM)) залежить від виду подання інформації – кількісної, якісної, релейної («так»/«ні»)) [6, 7], що є значним обмеженням їх застосування, так як у практичних завданнях, зокрема при розв’язанні задачі прийняття рішень щодо вибору «дисциплін за вибором», інформація може бути представлена різними типами даних. Інформація може бути подана у вигляді конкретних числових даних, бути статистичною, містити елементи випадковості або ж стосуватися збору та обробки думок людей, які мають власний погляд на проблему, що вирішується. Слабо формалізовані або надмірно інформаційні проблеми розв’язуються методами експертного оцінювання. Методи експертного оцінювання визначають шляхи підбору та аналізу, використовуючи досвід професіоналів, аналітиків і практиків, тобто людей із креативністю, конструктивним підходом і колективним мисленням у процесі оцінювання одних чи інших питань [8-14].

У роботі математичну основу становлять підходи, засновані на методах оптимізації, дослідження операцій та експертного оцінювання, які дозволяють здійснити кількісну оцінку та визначити пріоритетність кожної дисципліни стосовно інших із запропонованого списку.

При застосуванні методів оптимізації та дослідження операцій оптимальний варіант прийнятого рішення знаходиться з багатьох можливих. Процедура пошуку здійснюється шляхом порівняння варіантів. Варіант рішення, це складова одиниця. Безпосереднє їх порівняння неможливе. Кожен варіант оцінюється. Функціональна

оцінка має цільове призначення. Отже, оптимізація пов'язана з використанням трьох елементів: варіант, безліч варіантів, критерій [16]. У математичному уявленні це записується наступним чином, нехай x - варіант рішення, що складається із заданої кількості складових одиниць - компонент, кількісні значення яких представляють сам варіант; G - безліч можливих варіантів; функціональна оцінка варіанта - $f(x)$. Математичне подання вибору оптимального варіанта запишеться у вигляді деякої математичної моделі

$$Z = extr\ f(x), \quad x \in G. \tag{1}$$

Проблема полягає у підборі самого варіанта (набору компонент), математичного опису безлічі варіантів та визначенні аналітичного виду функції – оцінки варіанта.

Метод пошуку оптимального варіанта розв’язання залежить від усіх елементів оптимізації та їх математичного представлення, тобто від математичної моделі

Основним математичним апаратом розв’язання оптимізаційних задач є математичне програмування (МП). Формально задача математичного програмування каже: необхідно визначити екстремум багатовимірної функції $f(x_1, x_2, ..., x_n)$ на опуклому багатограннику, що задається сукупністю інших багатовимірних функцій $g_i(x_1, x_2, ...x_n)$ $(i = \overline{1,m})$. Додаткові вимоги накладаються на аргументи всіх цих функцій, тобто на x_j $(j = \overline{1,n})$ Математична модель задачі МП має вигляд

$$Z = extr\, f(x_1, x_2, ...x_n) \tag{2}$$

при обмеженнях

$$g_i(x_1, x_2, ...x_n) \le 0 \quad (i = \overline{1,m}) \tag{3}$$

та додаткові вимоги

$$x_j \in D_j \quad (j = \overline{1,n}). \tag{4}$$

Функція $f(x_1, x_2, ..., x_n)$ називається цільовою функцією, що відображає критерій оптимізації. Система обмежень - математичний опис безлічі варіантів, що визначає, які з них є допустимими. Сам варіант є сукупністю x_j, що представляється у вигляді вектору $X = (x_1, x_2, ..., x_n)$,компонентами якого можуть бути числа, які мають певну характеристику D_j.

Метод розв’язання задачі МП це перебір варіантів – векторів, $X^k = (x_1^k, x_2^k, ..., x_n^k)$, $k = 1, 2, ...$, що задовольняють обмеженням і містять компоненти x_j^k заданої характеристики D_j. Метою перебору є визначення такого варіанту $X^* = (x_1^*, x_2^*, ..., x_n^*)$, який приведе цільову функцію до екстремального значення Z^* (5).

$$Z^* = f(x_1^*, x_2^*, ..., x_n^*). \tag{5}$$

Процедурно метод розв'язання задач МП залежить від: функцій, що визначають модель; сутності додаткових вимог D_j; характеристики змінних, що входять до запису $f()$ та $g_j()(i = \overline{1,m})$. Конкретизація всіх трьох аспектів моделі завдання МП визначає підхід і сам метод розв'язання. Задачі МП бувають лінійними, детермінованими, а також стохастичними. Різні класи задач наведено на рис.1.

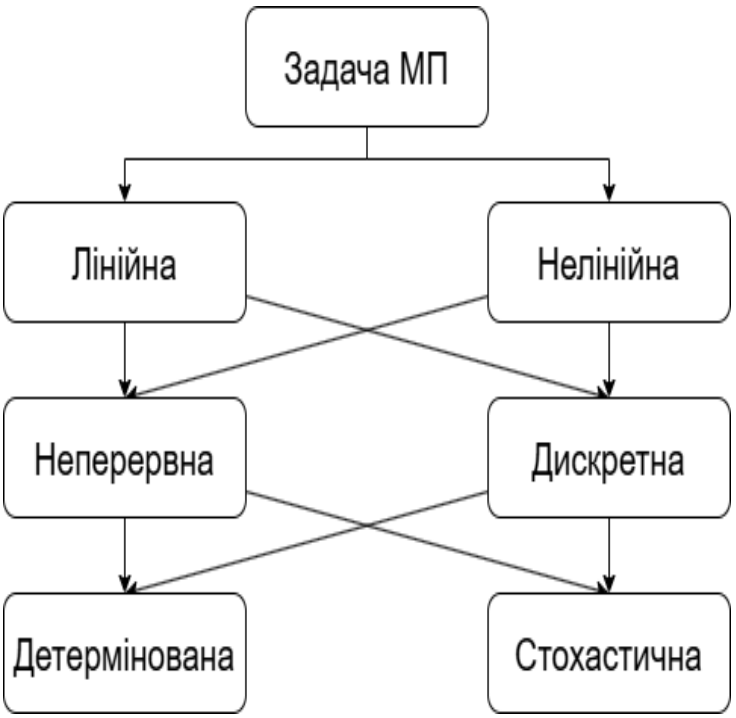


Рис.1. Класи задач МП

У обчислювальному сенсі моделі задачі МП бувають з кінцевою та нескінченною кількістю варіантів. Від цього залежить підхід та ідея методу реалізації завдання кожного класу. Спрямований перебір варіантів реалізації моделей нескінченною кількістю варіантів організується на основі так званих j-их аналітичних методів, розроблених з використанням певних прийомів чисельних методик. Другий підхід комбінаторний, заснований та використовуваний для задач з кінцевою кількістю варіантів. Комбінаторні методи мають перестановний характер. Вони націлені на ідею зменшення кількості варіантів, що перебираються. Обчислювальні процедури комбінаторних методів, це пошук частин безлічі варіантів, які свідомо не містять оптимального і підлягають відсіюванню.

Задача МП належить до класу детермінованих завдань лінійного програмування (ЛП), якщо її модель записується наступним чином:

$$Z = \max \sum_{j=1}^{n} c_j x_j \tag{6}$$

при обмеженнях

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = \overline{1,m} \tag{7}$$

та додаткових вимог до компонентів вектора рішень (варіанту рішень)

$$x_j \ge 0, \quad j = \overline{1,n}. \tag{8}$$

Додаткові вимоги роблять задачу безперервною, що зумовлює нескінченну кількість варіантів розв'язання. Детермінованість розуміється за умовчанням. Якщо додаткові вимоги доповнити фразою «x_j - ціле», то задача називається задачею цілісного програмування (ЦЛП), що відноситься до класу задач дискретного програмування і має кінцеву кількість варіантів.

Окремий випадок задач ЦЛП складають завдання ЛП з булевими змінними. Тоді задача має комбінаторний характер. Додаткові вимоги записуються як $x_j \in \{0,1\}, \quad j=\overline{1,n}$ і визначають підклас, в який включаються задачі про ранець [17, 18]. Особливістю цих задач є те, що змінні c_j, b_i, a_{ij} є позитивними числами. Задачі про ранець мають окремі методи розв'язання.

В роботі розглядається задача оптимального вибору дисциплін за вибором зі списку запропонованих, яка відноситься до класу задач лінійного програмування дискретного типу. Задача у прикладному сенсі відноситься до задачі з неподільностями, формалізується як завдання лінійного програмування з булевими змінними, вирішується комбінаторним методом і призводить до оптимального вибору по відношенню до специфічного критерію корисності підвищення компетентності майбутніх фахівців з вищою освітою з прикладної математики.

Експертне оцінювання. Дисципліни в навчальному плані, що йдуть під назвою «на вибір» у прямому сенсі не мають вирішального слова при підготовці фахівців. Проте вони націлені на підвищення компетентності і практичності претендентів за фахом. Проблема оптимального вибору насамперед пов'язані з підбором критерію оцінки якісної підготовки. Переліку такого типу критеріїв немає. Пропозиції варто отримувати насамперед від фахівців, які займаються підготовкою майбутніх спеціалістів, а також від тих, хто працевлаштовує та очікує сучасного, компетентного і конструктивного працівника. Загалом важливо враховувати думку тих, хто прагне бути корисним і впевненим членом колективу. У роботі пропонується використати методи експертного оцінювання. Збір та аналіз думок людей, які глибоко розуміють сучасні виклики вищої освіти, дозволить кількісно оцінити кожну дисципліну зі списку та визначити їх пріоритетність у виборі. Методи експертного оцінювання дозволяють отримувати кількісні оцінки якісних характеристик, які важко формалізувати. Вони застосовуються у проблемних областях із недостатньою структурованістю, а також у випадках надлишку інформації, коли потрібно виділити найбільш інформативні дані.

Експертне оцінювання, це багатоетапний процес, на кожному з яких приймається рішення, що забезпечує істинність результату. У першу чергу – визначення цілей експертизи. Цілі мають бути сформульовані так, щоб бути зрозумілими для всіх, хто бере участь у виконанні роботи під час оцінювання. Водночас вони можуть містити певні приховані аспекти, що дають змогу включати в процес особисте сприйняття розв'язуваної проблеми.

Підбір групи експертів і визначення її кількісного складу є непростим етапом у процесі отримання об’єктивних думок, на основі яких формується результат. Креативність експертів та їхня конструктивність у висловлюванні думок дозволяють суттєво розкрити суть проблеми. Важливими є також вміння працювати в колективі, сприймати думки колег, але при цьому не піддаватися тиску окремих учасників. Експерт повинен бути зацікавлений у прийнятті рішення, однак ця зацікавленість не повинна мати особистого характеру. Бажано, щоб ставлення до виконуваної роботи супроводжувалося максимальною віддачею. Саме такі вимоги пред’являються до потенційних експертів. Експерт не лише розуміє важливість розв'язуваної проблеми, але й відчуває її, інтуїтивно приймає рішення, не піддається впливу авторитетів, має власну думку та вміє її чітко висловити.

Кількісний склад групи експертів залежить від багатьох факторів. Наприклад, від термінів отримання результатів, фінансових можливостей організації, обсягів виконуваних робіт, географічного розташування експертів тощо. У математичній термінології - необхідно наявність ступенів свободи під час прийняття рішень.

Оптимальної кількості експертів у групі поки не встановлено. Один із можливих випадків, це встановлення деяких граничних оцінок $N_{\min}, N_{\max}$. Нижня границя безпосередньо залежить від кількості об'єктів, що оцінюються. Це один із найчастіше використовуваних величин під час підбору групи експертів. Вважається, що кількість експертів, які розглядають об'єкти, відповідає їх кількості. $N_{\max}$ розглядається як потенційно можлива величина.

Найбільш раціональний підхід – визначення числа експертів у групі, яка виконує оцінку, використовуючи методи статистичної обробки випадкових величин. Використовуючи коефіцієнт Стьюдента парі заданій величині достовірності результату, величина варіації (відхилення) повинна не перевищувати значення з інтервалу [0,2; 0,3].

Група експертів, це команда висококваліфікованих фахівців, які залучаються для вирішення проблеми. Однак, у першу чергу, необхідно мати список людей, серед яких обиратимуться експерти. Підбір людей-претендентів на участь в експертному оцінюванні залежить від важливості та рівня ієрархічного ланцюжка, в якому передбачається проведення експертного оцінювання. Якщо це глобальні проблеми країни, то експерт повинен мати статус міжнародного значення. Бути учасником експертного оцінювання вирішення проблем рівня міжнародних альянсів, знати задачі та способи їх вирішення на цьому рівні. Якщо проблеми мають більш прикладний характер, то їх вирішення не обов’язково залежить від міжнародних завдань і викликів. У цьому випадку група експертів складається з людей, які досконало розуміють специфіку ситуації та умови, в яких вони братимуть участь у процесі ухвалення рішення.

Для складання списків претендентів іноді використовують метод Шара (снігового кома), коли один із організаторів експертного оцінювання запрошує свого "знайомого", який, у свою чергу, запрошує свого знайомого і так далі. Таким чином, список поступово "набирає обертів", збільшуючи кількість осіб – претендентів для експертизи.

Інший підхід – самооцінка професіоналів, які бажають стати експертами. Організатори експертизи складають тестовий набір питань, кожне з яких має кількісну оцінку. Претендент, не знаючи цих оцінок, відповідає на поставлені питання. Якщо набрані бали не менше заданого порогу, що визначає право стати експертом, претендент зараховується.

Якщо експертиза проводиться в конкретному закладі, то проблема складання списку претендентів не виникає. Відомо, хто саме і для якої конкретної мети складається список претендентів або навіть експертів для прийняття рішень місцевого значення.

Результатом експертизи є узагальнена оцінка думок групи експертів. Достовірність отриманої інформації залежить від рівня професіоналізму всіх членів групи. Важливість думки експерта оцінюється кількісно і називається коефіцієнтом компетентності (КК). КК враховується при визначенні узагальненої оцінки.

Існують різні підходи до визначення коефіцієнта компетентності (КК) експертів. Найпростіший з них, який називається взаємооцінкою, визначається самими експертами. Кожен експерт групи присвоює число «1» всім іншим колегам, яких вважає «достойними» брати участь в експертному оцінюванні. КК визначається як відносна величина набраних одиниць до їх загальної кількості.

Оскільки експертне оцінювання як спосіб отримання інформації існує досить давно, було зібрано багато методів кількісної оцінки компетентності професіоналів, які

беруть участь у процесі прийняття рішень за питаннями експертного оцінювання. Можна назвати деякі з них. Досить популярним є тестовий метод отримання кількісних значень КК. Суть цього методу полягає в наступному: надається тестова таблиця оцінки професіоналізму експерта. Стандартним чином оцінюється кожен рівень перерахованих факторів. У таблиці наводиться числова величина P_{ij} ($i=\overline{1,m}$ – перелік факторів; $j=\overline{1,n}$ - рівень), що передбачає певну частку компетентності. Очевидно, що $\sum_{j=1}^{n} P_{ij}=1$ для будь-якого фактора i. Кожен конкретний експерт l відповідає одному з рівнів j по кожному конкретному фактору *i*, що позначимо через через $P_{ij_i}^{l}$. Це кількісна оцінка експерта l за відповідним фактором *i*. Величина $\sum_{i=1}^{m} P_{ij_i}^{l}$ це сумарна оцінка в балах, яку зібрав експерт l згідно представленої таблиці. Коефіцієнт компетентності експерта q_l визначається за формулою (9):

$$q_l=\sum_{i=1}^{m} P_{P_{ij_i}}^{l} \Big/ \sum_{\nu=1}^{m}\sum_{l=1}^{k} P_{ij_i}^{l} \quad l=1,2,...,k \tag{9}$$

де k – кількість експертів у групі.

Наведені дані кількісної оцінки факторів P_{ij} є коригувальними параметрами при баловій оцінці компетентності експерта. Згідно з цілями та завданнями експертного оцінювання ці параметри можна підбирати. Таким чином, коригувальні фактори дозволяють коригувати коефіцієнт компетентності (КК) експерта.

Сам підбір експертів і кількісна оцінка їх компетентності в першу чергу залежить від суті розв'язуваної проблеми. Якщо йдеться про організаційні питання, то краще за все проблему зрозуміють керівні працівники, які працюють у цій сфері. Технічні питання, їх складність і можливості вирішення можуть оцінити інженерні фахівці. Наукові питання мають вирішувати ті, хто володіє новими підходами, методами та техніками розв'язання в відповідній предметній області. Оцінити можна ту думку, яка в принципі є в наявності.

Експертне оцінювання як спосіб отримання кількісної інформації побудоване на основі «я так думаю» і організоване з урахуванням певних наукових підходів. В цілому це не статистика, отримана як результат спостережень за реальними подіями, це оцінка та обробка, з використанням методів статистичної обробки випадкових процесів. Дані, отримані методами експертного оцінювання, спочатку не є кількісними. Це результат того, як запитати та як сформулювати питання. Тому оцінити істинність, достовірність і точність результатів не так просто, а надати їм наукову основу досить складно. Оцінка достовірності залежить від узгодженості думок експертів. Існує коефіцієнт конкордації, який визначає ступінь розбіжності думок тих, хто оцінює проблему, на скільки їхні думки відображають істинність того, що відбувається. Це дуже важливо, якщо результат експертизи використовується для точних рішень, що передбачають глобальні наслідки. Якщо дані для прийняття поточного рішення і потрібна кількісна оцінка певних процесів чи об'єктів, то до отриманих результатів можна ставитися більш «м’яко». У такому разі можна задовольнитися тим, що одне і друге можна оцінити в кількісному вигляді і включити в процедуру обчислень.

Принципи розв’язання задачі. Задача, яка вирішується в даній роботі, стосується проблеми вибірки, критерій якої необхідно представити кількісно. Пропонується загальний список дисциплін із n позицій $(j=\overline{1,n})$. Серед них необхідно вибрати

m дисциплін ($m < n$). Вибір здійснюється з метою підвищення якості отриманої освіти, а саме - підвищення компетентності, практичних навичок та знань, необхідних для майбутньої професійної діяльності. По кожній пропонованій дисципліні виділяється t_j години навчального навантаження. Позначимо через c_j величину корисності j - ї дисципліни, яку можна отримати, вивчаючи j - у дисципліну зі списку. Необхідно вибрати дисциплін, яка забезпечить найбільшу сумарну корисність навчальної підготовки.

Математична модель оптимального вибору дисциплін зводиться до відомої математичної задачі, що носить назву «задача про ранець». Якщо в якості шуканих величин ввести булеві змінні такі, що

$$x_j = \begin{cases} 1, \text{ якщо } j-\text{а дисципліна додається до списку обраних}, \\ 0, \text{ в іншому випадку}. \end{cases} \quad (10)$$

Математична модель запишеться так (11-13):

$$Z = \max \sum_{j=1}^{n} c_j x_j \quad (11)$$

при обмеженнях

$$\sum_{j=1}^{n} t_j x_j \le A \quad (12)$$

та додаткових вимог

$$x_j \in \{0,1\},\ j = \overline{1,n}. \quad (13)$$

Реалізація моделі можлива при відомих n, t_j, A, c_j. Величина навчального навантаження з кожної дисципліни t_j відома, відомий список дисциплін – величина n і кількість годин, відповідальних у розділі освітньої програми «дисципліни на вибір» - величина A. Корисність кожної дисципліни c_j визначається експертним шляхом, оскільки без думок висококваліфікованих спеціалістів неможливо однозначно оцінити її значущість і користь для навчального процесу.

Тестовий приклад. Алгоритм визначення оптимального набору дисциплін за вибором продемонструємо на тестовому прикладі. З метою визначення корисності кожної дисципліни передбачається група експертів, яка включає людей, які представляють навчальний процес, сферу виробництва та студентства. Наприклад, професор кафедри, яка випускає, гарант спеціальності, куратор групи, методист, стейкхолдер, випускник, що працює на об'єкті, староста групи та активний студент групи. Позначимо експертів через $e_1, e_2, ..., e_8$. Оцінка думок експертів може проводитись за двома основними факторами: займане соціальне становище та науково-технічні досягнення. Далі розглянемо, як можна поділити групу експертів за цими факторами. Як приклад, припустимо, що по першому фактору групу експертів можна поділити на три підгрупи, а за другим на чотири. Нехай експерти e_1, e_3, e_5 складають першу підгрупу, e_2, e_4, e_6 - другу, а експерти e_7, e_8 - третю. Згідно з приналежністю до відповідної підгрупи експертам приписуються бали з інтервалу [1,3]. Припускаємо, що за другим фактором експерти діляться так: e_1, e_5 складають першу підгрупу, e_3, e_4 - другу, до третьої групи належать експерти e_5, e_6, e_7, а остання підгрупа представлена експертом e_8. Відповідно, система балів з чисел інтервалу [1,4]. Тестові дані наведені у таблиці 1.

Таблиця 1.

Тестові дані

Експерт		e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	Сума
фактор	1	3	1	3	1	3	1	2	2	16
	2	4	3	2	2	4	3	2	1	21
Сума балів		7	4	5	3	7	4	4	3	37
Коефіцієнт компетентності q_i		0,188	0,108	0,135	0,082	0,188	0,108	0,108	0,082	0,999

Кількість дисциплін згідно зі списком передбачуваних дорівнює 10 (n =10). Позначимо їх через $a_1, a_{2,...,} a_{10}$. Проведено опитування експертів. Отримані ранжирування. Умовні результати наведено у таблиці 2.

Таблиця 2.

Результати ранжування

Дисципліна / Експерт	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
e_1	1	6	4	2	3	8	5	7	10	9
e_2	2	3	4	1	5	7	6	9	8	10
e_3	3	4	2	5	1	9	7	6	10	8
e_4	2	4	3	5	1	6	7	8	9	10
e_5	1	2	4	3	5	7	6	9	10	8
e_6	1	4	5	2	3	10	9	8	6	7
e_7	2	1	3	4	5	7	6	9	10	8
e_8	4	2	1	5	3	8	7	6	10	9

Отримані ранги дисциплін переводяться в бали. Використовується така інтервальна шкала балів: за перше місце в ранжуванні надається десять балів, за друге — дев'ять і т.д. Кількість балів для кожної дисципліни обчислюється з урахуванням коефіцієнта компетентності експерта, який визначив місце дисципліни в ранжуванні. Дані наведені в таблиці 3.

Таблиця 3.

Ранги дисциплін в балах

Показник	Дисципліна										Сума
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	
Бали	8,916	7,600	7,591	7,798	7,798	3,214	4,560	3,215	1,729	2,455	54,97
Корисність	0,166	0,138	0,137	0,140	0,140	0,050	0,083	0,089	0,031	0,043	0,997
Перевага	1	4	5	2-3	2-3	8	7	6	10	9	

Найбільшу сумарну корисність навчального навантаження визначимо, виходячи з наступних тестових даних. Величина A, загальна кількість кредитів, виділених для вивчення дисциплін за вибором, становить 25 кредитів. З метою спрощення розрахунків значення корисностей c_j переведемо в іншу шкалу вимірювання, тобто помножимо на 1000. Дані наведені в таблиці 4.

Таблиця 4.

Представлені дані за новою шкалою вимірювання

Дисципліна	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Корисність c_j	166	138	137	140	140	50	83	89	31	43
Обсяг у кредитах t_j	4,5	4,0	4,0	4,5	4,0	4,5	4,0	4,0	3,0	3,0

Пріоритетна черга включення дисциплін до списку обраних визначається за алгоритмом Данцига [19]. Алгоритм визначає отримання оптимального рішення відповідної нецілочисельної задачі про ранець. Для цього визначається частка c_j / t_j та впорядковується в порядку не зростання. Кількість дисциплін, включених до списку, регулюється величиною A . Розрахункові дані наведено у наступній таблиці 5.

Таблиця 5.

Розрахункові дані

Дисципліна	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
λ_j	37	36	36	31	35	11	20	22	10	14
Обрані дисципліні	1	1	1	1	1	0	0	1	0	0

Оптимальне рішення згідно з обраним критерієм корисності представляє перелік з 6 дисциплін, які входять до числа вибраних. Вони визначають максимальну кількість корисності при підвищенні компетентності випускника. Слід зазначити, що складання списку можна здійснити на основі узагальнених оцінок експертів, результатом яких є пріоритетна черга дисциплін. На основі черги можна регулювати кількість вибраних дисциплін, якщо ввести поріг вибору.

Висновки. У роботі вибір навчальних дисциплін із запропонованого списку розглядається як задача оптимізації, що моделюється через задачу про ранець, одну з основних задач цілочисельного лінійного програмування. Умови вибору дисциплін визначаються за допомогою обмежень на навчальний час та обсяг викладання дисциплін. Критерій оптимізації має характеристики проблеми, що важко формалізуються, тому для кількісної оцінки дисциплін було застосовано метод експертного оцінювання. Для ілюстрації роботи методу надано приклад вибору дисциплін на основі тестових даних.

Дослідження показало потенціал математичних методів для оптимізації процесу вибору навчальних дисциплін, що може сприяти підвищенню компетентності здобувачів вищих навчальних закладів. Наданий короткий опис лінійних моделей передбачав вибір моделі задачі про ранець. Відсутність числових даних для оцінки якості розглянутих дисциплін компенсована експертним оцінюванням. Наведено попередній опис основних процедур експертної оцінки та можливостей їх інтеграції в процес формування навчальних планів.

Список літератури

1. Лебедик Л. В., Стрельніков В. Ю., Стрельніков М. В. Сучасні технології навчання і методики викладання дисциплін : Навчально-методичний посібник для слухачів курсів підвищення кваліфікації педагогічних працівників закладів середньої, професійної (професійно-технічної), фахової передвищої та вищої освіти. Полтава: АСМІ, 2020. 303 с.

2. Kozina Y.Y., Verbitskaya E.V. Analysis of decision making methods to include a set of disciplines in the educational process. *Proceedings of the International Scientific Conference Intellectual Systems of Decision Making and Problems of Computational Intelligence ISDMCI.* 2019. Ukraine. P. 81–82.
3. Ілійчук Л. В. Сучасні стандарти та показники оцінювання якості вищої освіти в Україні. *Scientific Journal of Khortytsia National Academy*. 2022. Т. 2(7). С. 47-59. URL: https://doi.org/10.51706/2707-3076-2022-7-5
4. Рафальська О. О. Алгоритми формування базової таксономії дисциплін. Управління розвитком складних систем. 2015. №24. С.137-141.
5. Kostoglou V., Kafkas K. Design and development of an interactive mobile-based decision support system for selecting higher education studies. *Proceeding of the 8th Balkan Region Conference on Engineering and Business Education. Sibiu, Romania*. 2017. P. 240–248. DOI: 10.1515/cplbu-2017-0032
6. Kozina Y., Volkova N., Horpenko D. Mobile decision support system to take into account qualitative estimation by the criteria. *IEEE Third International Conference on Data Stream Mining & Processing (DSMP).* 2020. P. 357-361). URL: https://doi.org/10.1109/DSMP47368.2020.9204134.
7. Horpenko D. R. A conceptual model of decision-making support of the volunteer team in conditions of dynamic changes. *Herald of Advanced Information Technology*. 2022. V. 5. No. 4, P.275 – 286. DOI: 10.15276/hait.05.2022.20.
8. Архипов О. Є., Архіпова С. А. (). Оцінювання якості роботи експертів за даними багатооб'єктної експертизи. *Ukrainian Information Security Research Journal*. 2011. URL: https://doi.org/10.18372/2410-7840.13.2049
9. Величко О.М., Коломієць Л.В., Гордієнко Т.Б., Шевцов А.Г., Карпенко С.Р., Габер А.А. Групове експертне оцінювання та компетентність експертів. Одеса: ВМВ, 2015. 286 с.
10. Грабовецький, Б. Є. Методи експертних оцінок: теорія, методологія, напрямки використання: монографія. Вінниця: ВНтУ, 2010. 171 с.
11. Петренко Л.М. Педагогічна експертиза: Технологія експертного оцінювання результатів навчальних досягнень учнів. Харків: Основа, 2007. 176 с.
12. Подолянчук С. В. Визначення компетентності експертів з оцінювання наукової діяльності у вищому педагогічному навчальному закладі. К., 2015. 122 с.
13. Данченко О. Б., Царик Т. Ю., Савельєва Т. В., Занора В. О. (). Метод експертної оцінки ризиків навчального процесу в умовах модульно-рейтингової системи навчання. *Управління розвитком складних систем.* 2013. №13. С.143–146.
14. Ярощук Л. Експертні методи в автоматизованих системах керування. Формування та напрями використання експертних знань. Київ : КПІ ім. Ігоря Сікорського, 2022. 43 с.
15. Григорків В.С., Григорків М.В., Ярошенко О.І. Оптимізаційні методи та моделі. Чернівці : Чернівецький нац. ун-т, 2024. 464 с.
16. Юхименко Б. И., Гуляева Н. А. Методы оптимизации и исследование операций. Учебное пособие по самостоятельной работе. Одесса: Феникс, 2018. –204 с.
17. Васянин В., Востров Г., Вычужанин В. Информационные управляюшие системы и технологии, проблемы и решения, Одесса: Экология, 2019. С. 211–225. URL: http://dspace.opu.ua/jspui/bitstream/123456789/10153/1/Inner_2.pdf
18. Юхименко Б. И., Волкова Н. П. Наближені алгоритми розв'язання задачі про багатовимірний ранець. Дослідження в математиці і механіці, 2017. V.22(2). P. 105-116. URL: http://rmm-journal.onu.edu.ua/article/download/135745/pdf_

Б. І. Юхименко

MATHEMATICAL METHODS FOR OPTIMAL SELECTION OF ELECTIVE COURSES IN HIGHER EDUCATION INSTITUTIONS

B. I. Yukhymenko

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: biruteyu@gmail.com

The work is dedicated to the research and application of mathematical methods for the optimal selection of elective courses in educational programs at higher education institutions. In the context of general computerization, the processes of training specialists require the implementation of innovative approaches that ensure the increase of competence and practicality of future professionals. The methodological aspects of mathematical programming and expert evaluation, which are used for quantitative analysis and selection of the most priority courses, are discussed. Special attention is given to the task of optimally choosing elective courses from a list of proposed ones, based on criteria of utility and academic workload. The problem is formalized as a discrete-type linear programming problem, known as the "knapsack problem." The work proposes an approach based on expert assessments to determine the utility of courses, as well as the construction of mathematical models and methods for solving them. The research emphasizes the importance of integrating expert evaluation methods into the process of curriculum formation, which contributes to improving the quality of education and preparing professionals who meet the requirements of the modern economy and society. The work holds practical value for educational institutions and curriculum developers focused on labor market needs.
Keywords: Expert evaluation, linear programming, courses, educational programs, knapsack problem.

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

Том 14, номер 4, 2024. Одеса – 418 с., іл.

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Volume 14, No. 4, 2024. Odesa – 418 p.

Засновник: Національний університет «Одеська політехніка»

Зареєстровано Міністерством юстиції України 04.04.2011р.
Свідоцтво: серія КВ № 17610 - 6460Р

Друкується за рішенням Вченої ради Національного університету «Одеська політехніка», (протокол №5 від 24.12.2024р.)

Адреса редакції: Національний університет «Одеська політехніка»,

1, Шевченка проспект, Одеса 65044 Україна

Web: www.immm.op.edu.ua (immm.opu.ua)

Email: immm.ukraine@gmail.com