

ARCHITECTURE AND TRAINING ALGORITHM FOR NEURAL NETWORK TO RECOGNIZE VOICE SIGNALS

Molchanova V. S. – PhD, Associate Professor, Department of Informatics, SHEU “Pryazovskiy State Technical University”, Mariupol, Ukraine.

Mironenko D. S. – PhD, Head of department, Department of Informatics, SHEU “Pryazovskiy State Technical University”, Mariupol, Ukraine.

ABSTRACT

Context. Typically, interaction between user and mobile devices is realized by touchings. However, many situations, when to implement such interaction is too awkward or impossible, exist. For example, with some diseases of musculoskeletal system, motility of movements may be impaired. It leads to inability to use device efficiently. In that case, a task of looking for alternative ways of person-device interaction becomes relevant. Voice interface development can be one of the most prospective tasks in that way.

Objective. The goal of the study is to develop a project of neural network architecture and internal components for voice-controlled systems. Resulting interface have to be adapted for processing and recognition Ukrainian speech.

Method. An approach, based on audio signal analyzing by sound wave shape and spectrogram, is used for making got via microphone data, appropriate for processing. Using neural network makes possible sounds classification by generated audio signal and information of its transcription. The neural network structure is completely adapted to peculiarities of Ukrainian phonetics. It takes into account the nature of the sound wave, generated during sound pronunciation, as well the number of sounds in Ukrainian phonetics.

Results. Experiments were carried out aimed to choosing optimal neural network architecture and training sample dimension. The root-mean-square deviation of neural network error was used as the main criterion in assessing its effectiveness. A comparative analysis of effectiveness of the proposed neural network and existed on the market speech recognition tools showed improvement in the relative measures of recognition by 9.26%.

Conclusions. Obtained in the research results can be used for full-featured voice interface implementation. Despite the fact that the work is focused on recognition Ukrainian speech, the proposed ideas can be used during developing transcribing services for other languages.

KEYWORDS: voice interface, audio signal, signal amplitude, spectrogram, neural network, training set, standard deviation.

NOMENCLATURE

$A(t)$ is a dependence of sound signal amplitude from time over a continuous time period;

$A'(t_k)$ are discrete values of audio signal amplitude;

e_i is an experimental error of learning a neural network a i -th iteration;

E is a permissible error of neural network learning;

f_j is a function of converting the sound wave characteristics at definite time moment to a sound reference, which generates sound wave with corresponding characteristics;

h is a neuron of a hidden layer of a neural network;

H is a set of hidden layer neurons;

$H(t)$ is a dependence of sound signal frequency from time during some continual time period;

$H'(t_k)$ are discrete values of sound wave frequencies;

i is a neuron of an input layer of a neural network;

I is a set of hidden layer neurons;

J is an index of a word, sound, or speech fragment in a predefined alphabet R ;

L is a size of training sample;

n is a size of output alphabet;

N is a number of neurons in the neural network;

N_I is a number of neurons in an input layer of the neural network;

N_H is a number of neurons in a hidden layer of the neural network;

N_O is a number of neurons in an output layer of the neural network;

o is a neuron of an output layer of a neural network;

O is a set of output layer neurons;

P is a training sample;

p_j is probability of matching the sound wave to an element;

R is an alphabet of sounds and words;

r_j is a single, predefined sound, word, or speech fragment;

t is a time;

t_k is a discrete time;

T is a test sample.

INTRODUCTION

Voice recognition methods in conjunction with speech-to-text technologies is a very important tool for creating voice interfaces. Quality voice interface can act as a full-fledged alternative to traditional interactive one for hands-free systems. Such services are especially useful for users, suffering from musculoskeletal disorders. Due to impaired movements coordination, it is difficult and often almost impossible for them to interact device touching screen. Other way for using voice interface is in situations when user's hands are busy with executing another action.

The object of research is a process of speech -to-text converting.

The subject of research is neural networks application in development transcription services adapted on pronunciation features in Ukrainian, able to recognize separate speech fragment.

The purpose of the paper is to develop a method for recognition voice commands and to implement based on it service, which provides a voice interface for devices running Android. As a main mathematic tool is supposed using neural networks. This service will let use phone, tablet, etc. fully, without necessity to touch its screen by fingers. The service will be represented as an additional add-in, which can recognize voice commands in natural human language and translate them into device control commands.

To achieve the goal, it is necessary to solve a number of tasks:

- 1) To analyze main characteristics of sound waves generated when sounds and words are pronounced in Ukrainian;
- 2) To consider existing approaches to recognizing sounds by nature of their sound waves from the viewpoint of possibility to adapt them for recognition voice commands in Ukrainian;
- 3) To develop architecture of the neural network and to carry out its training;
- 4) To test the proposed neural network and to compare its efficiency in recognizing voice commands in Ukrainian to existing voice transcription tools efficiency.

1 PROBLEM STATEMENT

Let some alphabet $R = \{r_1, r_2, \dots, r_n\}$, is given, each its element r_i can correspond to a separate sound, word or speech fragment.

As sound signal has arrived, sound wave is generated. According to [1–3], dependences of sound wave amplitude and frequency from time are main characteristics of sound wave which identify uniquely the nature of sound. It lets identify sounds basing on sound wave data and/or spectrogram for a certain time period (1, 2), as well as perform the inverse transformation digital data of the sound wave into analog sound.

$$f_1(A'(t_k), t_k) \rightarrow \{p_j\}_n, \quad (1)$$

$$f_2(A'(t_k), H'(t_k), t_k) \rightarrow \{p_j\}_n. \quad (2)$$

The functions f_1 and f_2 implement determination probability of correspondence signal, which has generated the investigated sound wave, and each sound from the alphabet R . The first one uses only its amplitude during a certain short time period, the second one has an extra parameter, which let use extra information about the signal frequency. The sound classification process is implemented by artificial neural networks.

The final decision, whether the sound signal corresponds to some element r_j is taken based on corresponding to it maximum $p_j(3)$:

$$r_j \rightarrow p_j : p_j = p_{\max} = \max(p_j), p_{\max} \gg 1 - p_{\max}. \quad (3)$$

Designing the set R , it is appropriate to include one element corresponds to an empty result, i.e. to a situation, when no one sound from set has been identified. Restriction $p_{\max} \gg 1 - p_{\max}$ at (3) let avoid false positives, when found probability maximum isn't quite large to classify sound generated sound wave to corresponding class.

2 REVIEW OF THE LITERATURE

The idea to develop a service, which is capable to recognize voice commands and speech in a natural language, isn't essentially new. Developments in speech recognition and voice controlling have been ongoing for over 20 years.

The Dragon Dictate Naturally Speaking system [4] was one of the first software systems capable to perform human speech recognition. In some cases, its recognition accuracy reached 95%. However, high recognition rates were achieved only for speech in English with certain pronunciation rate. In 1997, an attempt to adapt this system for recognition speech in Russian was made. Thus, the Gorynych system was developed [5]. The Gorynych system supported possibility to dictate text and to control some Windows functions with a voice. Meanwhile, speech recognition quality rarely exceeded 30%, which isn't an acceptable result. No attempt to adapt this system to speech recognition in Ukrainian has been made.

This direction of researches became especially popular when mobile devices running on Android and IOs operating systems, appeared. Nowadays, next voice assistants are the most popular: Siri (Apple) [6], Alexa (Amazon) [7, 8], Google Voice Assistant [8], etc. However, they are focused on voice input of text messages, which afterwards usually send via standard messengers, and keywords, which are used for searching information on the Internet. These tools provide satisfactory results within declared capabilities. Meanwhile, they are sensitive to pronunciation quality as well as to speech timbre. An essential disadvantage of these systems is necessity for permanent Internet connection. This disadvantage is caused by necessity of voice processing on the server side of application. In addition, in this case, whether confidentiality of the transmitted information will be retained and possible further ways of using it remain unclear. Another essential disadvantage of these services is limited set of available languages. So, for example, none of above mentioned voice assistants is able to recognize voice commands in Ukrainian. In addition, these systems are only voice assistants. None of them is a fully-featured voice interface.

The technologies of converting speech into a command or text used for developing voice assistants can also be used for creation full-fledged voice interfaces. Moreover, developers provide complete API for this: Google Speech API, YandexToolkit API etc. In addition, there are specialized platforms, which main task is to convert voice to text, for example, PocketSphinx. Development of an original service based on these technologies will provide a full-fledged voice interface for devices managed by

Android. However, it won't provide to fix the rest disadvantages of existing voice recognition services. Therefore, it is useful to develop an original application, which core will be focused on linguistic features of Ukrainian.

As a result of review literature sources devoted to the problem of converting speech into text [9–12], a number of tasks, which have to be solved sequentially were identified, as well as results which have to be obtained in course of solving each of them (Fig. 1).

At this stage the most interest represents the task of recognizing voice signals and converting them to text.

There are 3 main approaches to speech recognition algorithms implementation: hidden Markov models, dynamic programming, and artificial neural networks.

The approach based on hidden Markov models [13] needs long-term system debugging on large sets of test samples. This approach is quite simple from implementation viewpoint. In addition, enlargement set of recognizable words does only a little effect to computational complexity increase. However, it doesn't guarantee high accuracy of result, because to estimate error value reliably isn't always possible

The approach based on using dynamic programming [14] presupposes comparing two speech segments and determination difference indicator between them. Known in advance pattern is used as a first segment, identifiable – as a second. Using dynamic programming in this approach lets perform optimization and determine the template, which most accurately matches the recognized one. This approach lets get good results for low time and computational costs for small data samples, upon a condition, if a recognizable pattern matches to one element is in set. However, even slight increase of test data sample or outputs variants leads to significant complication of the calculation model.

The most powerful tool for solving speech recognition problem is artificial neural networks [15]. This approach provides not only individual words and sounds recognition, but continuous speech. Using neural networks represents the most interest for development speaker-independent speech recognition systems. Nevertheless, in a cause of complexity of neural network structure determining and its proper training, using neural networks

would be recommended only if two previous approaches proved ineffective.

Taking into account specifics of the project being developed (necessity for quick adaptation to specifics of each person's pronunciation and possible minor diction disturbances characteristic for disabled), approach based on using neural networks is the most prospective.

3 MATERIALS AND METHODS

Received from microphone sound signal is a sound wave with continually changing frequency $H(t)$ and amplitude $A(t)$. Amplitude determines the sound volume, and frequency – its tone. At the same time, digital sound processing and recognition can be performed only for discrete data sets $A'(t_k)$, $H'(t_k)$. Thus, for further full-fledged signal processing, it is necessary to perform transformations (4, 5):

$$A(t) \rightarrow A'(t_k). \quad (4)$$

$$H(t) \rightarrow H'(t_k). \quad (5)$$

The procedure to replace continuous dependence of sound characteristics to discrete ones, is called sampling, in doing so sampling frequency determines quality of result discrete signal. Usually the sampling rate is in range of 8–48 kHz.

As patterns were used discretized sound signals, gained in pronouncing all kinds of sounds and their typical combinations, characteristic for Ukrainian phonetic (Table 1).

Fig. 2 shows typical fragments of sound waves generated while some sounds of Ukrainian phonetic are being pronounced.

Shapes of sound waves, as well as spectrograms, make it possible to come up with information about amplitude $A'(t_k)$ and frequency $H'(t_k)$ of sound wave at each point of discrete time scale. This data can be used by the neural network to classify sounds.

Nowadays, multilayer neural networks are the most popular and efficient tool in speech recognition.

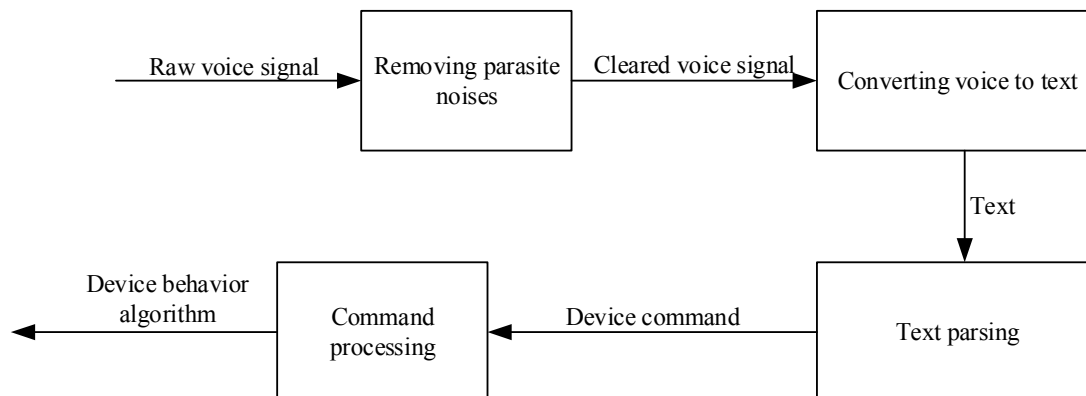


Figure 1 – A structured scheme of a typical voice command recognition and procession system

Table 1 – Transcriptions of some Ukrainian alphabet letters and its combinations

№	Letter	Transcription
1	А а	/a/, /a/
2	Б б	/b/
3	В в	/v/, /w/, /w/, /w/
4	Г г	/h/
5	Ґ ґ	/g/
6	Д д	/d/
7	Е е	/e/, /e/
8	Є є	/je/, /je/
9	Ж ж	/z/
...

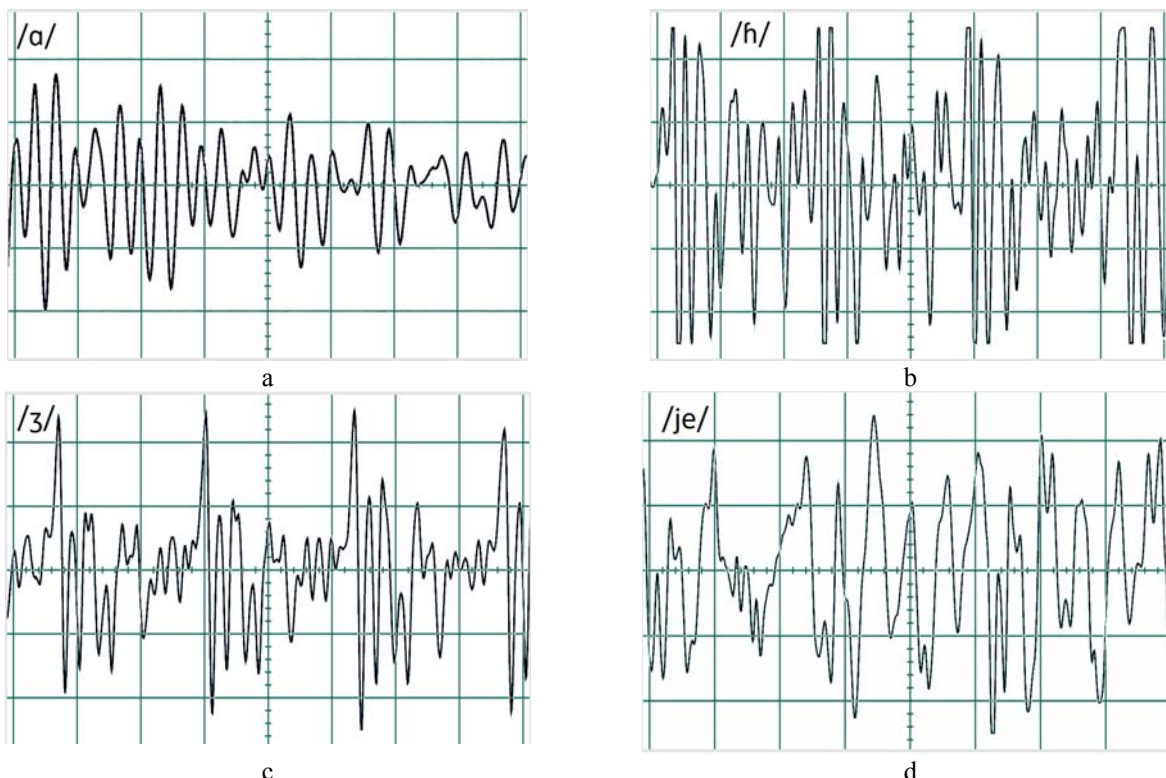


Figure 2 – Examples of sound waves shapes for some sounds of Ukrainian phonetic

To solve this task, it is assumed to use a three-layer neural network, which contains an input layer, one hidden layer and an output layer. Number of neurons are in input layer is defined as N_i , in output layer – N_o , in hidden layer – N_h . Denote by $I = \{i\}_{N_i}$, $O = \{o\}_{N_o}$, $H = \{h\}_{N_h}$ elements are in input, output, and hidden layers, respectively. Number of all neurons in the neural network is labeled as N , limit permissible learning error value as E .

Speech recognition will be performed based on the sound waveform data. To the input of neural network, data is provided sound wave frequency at each discrete time instant.

This approach implementation will lead to necessity to use input data vectors with length of 1000 or more elements. Algorithms used for the corresponding neural network learning will be demanding on computing resources,

and got neural network will not always be able to provide result required accuracy.

To optimize neural network structure, the assumption was made that it is appropriate to take in attention only extreme values of the sound waves amplitudes $A(t)$ and time moments t , when they're got detected (Fig. 3). Later, information about sound wave amplitudes between extreme values can be obtained by linear approximation.

Consequently, to inputs of the neural network, pairs $(t, A(t))$ will be supplied, and its number will be reduced to 20–30.

Number of output layer neurons corresponds to number of sounds which have to be recognized. Ukrainian phonetics involves 38 different sounds. Thus, number of neurons in the neural network output layer needed to recognizing every individual sound in Ukrainian phonetic is 38. As an activation function of the output layer a linear function is chosen.

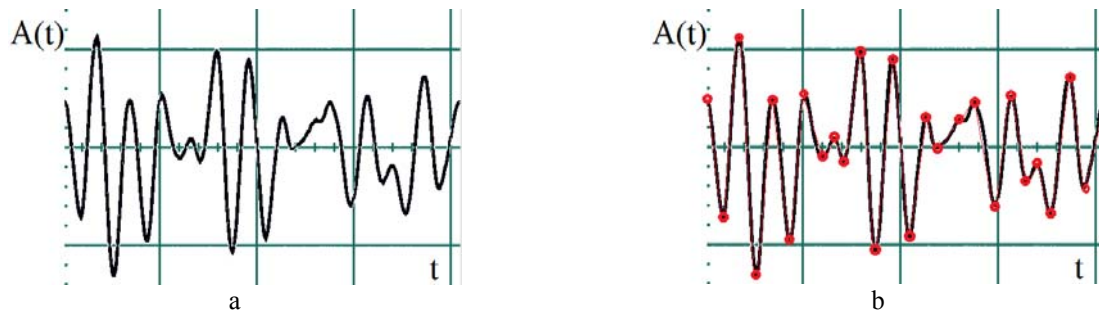


Figure 3 – Linear approximation of sound wave sections between it's extremums

Number of hidden layer neurons depends on some different factors: as training sample size, number of input layer neurons, number of output layer neurons, etc. An initial approximation of this value N_{H_0} can be defined as the average between number of input and output layers neurons. Afterwards a learning error e_i is calculated. Further optimization of neural network structure is performed by reducing (increasing) number hidden layer neurons and constructing a learning curve. The most optimal is considered solution, which provides learning error the closest to acceptable neural network learning error.

$$\begin{aligned} N_{H_0} &= \frac{N_i + N_O}{2}, e_0 = E(N_{H_0}), \\ N_{H_i} &= N_{H_{i-1}} \pm 1, e_i = E(N_{H_i}), \\ |e_i - e| &\rightarrow \min. \end{aligned} \quad (6)$$

As an activation function of hidden layer neurons hyperbolic tangent is chosen.

Created neural network was trained by the direct error propagation method. Each element in training sample P^L contains a vector, which dimension corresponds to number of neural network inputs, and a single integer value, which determines corresponding output of the neural network. Each element of training sample vector is represented as a pair of values $(A(t), t)$.

For three-layer neural network training sample size L is determined by relationship (7):

$$2 \cdot (N_i + N_H + N_O) \leq L \leq 10 \cdot (N_i + N_H + N_O). \quad (7)$$

Thus, a training sample P^L is formalized by the expression

$$P^L = \left\{ \left(\begin{array}{c} (A(t_1), t_1), (A(t_2), t_2), \dots \\ (A(t_{N_i}), t_{N_i}) \end{array} \right), o_k \in \{O\}_{N_0} \right\}_L. \quad (8)$$

Final assessment of developed neural network efficiency is executed on test samples T , which hasn't been used in neural network training. Test sample for the proposed in this paper neural network is formalized by expression (9)

$$T = \left(\left(\begin{array}{c} (A(t_1), t_1), (A(t_2), t_2), \dots \\ (A(t_{N_i}), t_{N_i}) \end{array} \right), o_k \in \{O\}_{N_0} \right). \quad (9)$$

It should be noted that during test process, the neural network uses ready output values only to estimate error, but not to improve result.

4 EXPERIMENTS

Experiments on the developed neural network were carried out using the original application Voicer, which implements the neural network proposed in the paper. The application is adapted for running on any devices managed by Android. Android Studio was chosen as the development environment, because of nowadays it is the most popular tool for developing Android applications. The neural network is implemented with TensorFlow Mobile library.

The application Voicer has a friendly user interface and let recognize the voice command received from the device's microphone without necessity to delve into the structure and principles of the neural network.

During the experiment was being carried out, various neural network architectures were analyzed, as well as options for number of samples in training set. As determining criterion for choosing neural network structure was admitted the standard deviation MSE. The module of difference between its value and permissible neural network error has to be minimal. It should be noted that we shouldn't try to minimize MSE to 0 value, in this case the effect of retraining neural network is possible and as a subsequent, incorrect result will be got on test samples.

5 RESULTS

Preliminary calculations have let limit number of neurons in the hidden layer (6) in range from 20 to 38 and training sample dimension (7) in range from $2 \times N$ to $10 \times N$. However, not each of possible architectures lets create a neural network efficient to solve assigned task. As an indicator of the neural network efficiency mean squared error was used. We've tested each of available architectures and estimated its efficiency. Results are assembled in table 2. Table columns contain dimension of test sample set, rows – number of hidden layer neurons and in the cells are mean square error values for respective neural network and test sample set.

Presented in Table 2 experimental results show that the best quality in recognizing sounds in Ukrainian was got if training sample set dimension was in range from $4 \times N$ to $8 \times N$, and number of hidden layer neurons was in range from 24 to 31. For illustrative purposes, this area is highlighted in the table and data, contained in, are presented in diagram (Fig. 4). Separately, minimum mean square error values for each number of elements in training sample set were highlighted.

Examined alternative tools for speech recognition and compared their efficiency to efficiency of developed one. To pursuit of testing process some commonly known voice assistants as Siri, GoogleVoice Assistant, Alexa were used. The initial sample of words is constructed in such a way that it uses all sounds characteristic of pronunciation in Ukrainian. Each word was pronounced 100 times by different voices, with different intonations and timbre. Information about number of correct recognitions is summarized in table 3.

Table 2 – Results of experiments

$L \backslash N_H$	$2 \times N$	$3 \times N$	$4 \times N$	$5 \times N$	$6 \times N$	$7 \times N$	$8 \times N$	$9 \times N$	$10 \times N$
20	7.39	4.56	1.90	10.27	15.4	15.78	16.34	18.23	23.98
21	5.62	4.21	1.76	9.81	12.1	15.26	17.39	20.35	25.35
22	4.62	3.76	0.52	7.92	8.87	14.83	15.82	13.50	18.76
23	3.28	5.12	0.31	3.96	5.12	13.67	10.05	5.52	13.34
24	1.40	1.22	0.08	2.39	4.63	4.36	4.52	7.84	12.18
25	5.06	5.57	0.26	1.59	1.89	7.12	5.65	12.96	14.73
26	5.09	3.58	0.28	1.12	1.77	5.65	5.95	13.14	18.23
27	6.28	5.75	0.25	1.28	0.91	2.09	3.39	12.96	20.34
28	7.25	2.85	0.54	0.69	0.15	0.77	1.34	12.67	14.36
29	8.45	2.58	1.24	0.92	0.28	0.98	1.73	8.93	12.34
30	10.63	5.84	2.83	2.63	0.81	0.24	3.56	7.65	15.28
31	14.48	7.36	3.25	2.09	1.95	0.56	7.82	12.84	18.56
32	12.54	4.56	5.60	4.89	3.12	0.33	4.23	9.23	12.90
33	15.94	12.34	8.96	6.98	5.86	1.98	7.12	34.19	18.76
34	16.75	15.11	11.7	12.03	12.98	2.09	15.74	20.65	21.63
35	20.73	18.39	12.1	15.09	19.27	4.51	10.09	23.12	23.45
36	23.86	20.00	11.3	18.82	21.43	10.94	9.33	8.95	28.32
37	23.12	21.98	14.0	16.31	20.60	12.25	17.52	12.51	20.52
38	15.94	16.93	14.2	18.59	21.85	23.15	25.12	23.18	18.74

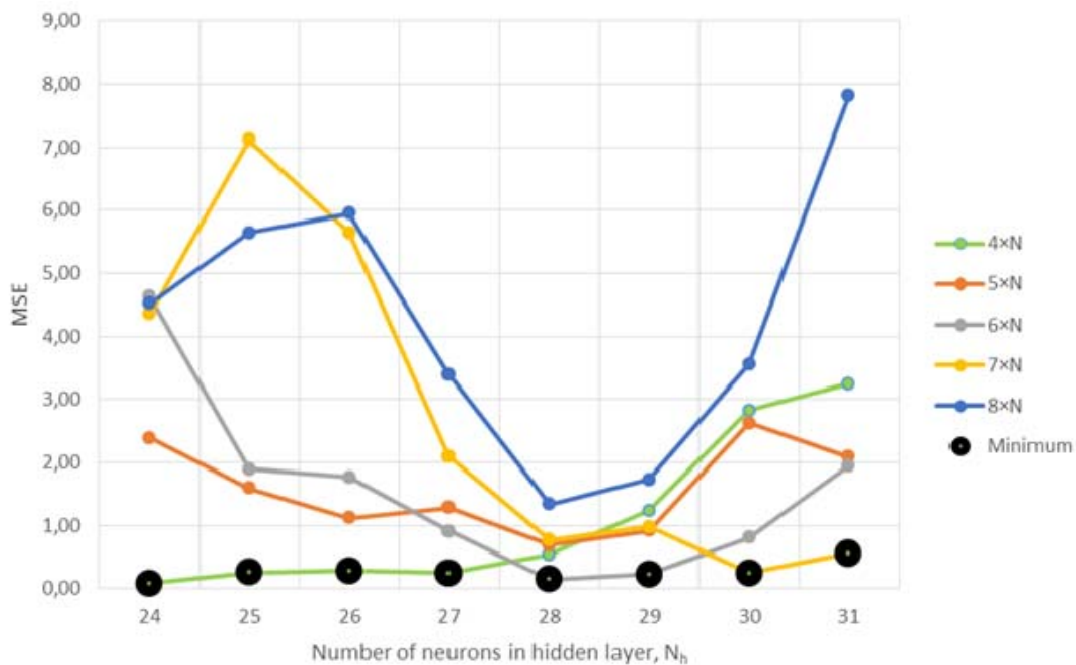


Figure 4 – Dependence of mean square error of the neural network from number of neurons are in hidden layer and dimension of training sample set

Table 3 – The number of correct recognition Ukrainian words got by various tools

Word	Transcription	Siri	Google Assistant	Voice	Alexa	Cortana	Proposed tool	Relative improvement in recognition quality
Дош	[до ш ч]	70	77		65	68	91	15.38%
М'який	[м й а к и й]	73	75		54	66	88	14.77%
Сьогодні	[с о г о д н і]	77	78		56	37	89	12.36%
Прикмета	[п р и к м е т а]	65	70		67	76	87	12.64%
Підсолоджувач	[п і д с о л о д ж у в а ч]	82	75		45	55	80	-2.50%
Мережа	[м е р е ж а]	77	90		57	46	90	0.00%
Відпустка	[в і д п у с т к а]	80	74		45	36	88	9.09%
Боротьба	[б о р о д " б а]	75	70		65	50	83	9.64%
Місяць	[м і с ' а ц ']	79	81		74	55	92	11.96%
Average		75.33	76.67		58.67	54.33	87.56	9.26

6 DISCUSSION

In accordance with table 2, we can conclude that there are several cases, when the best result is achieved. So, for example, for a neural network, which number of hidden layer neurons is in range from 24 to 27 neurons, optimal dimension of training sample is $4 \times N$. If number of hidden layer neurons is 28 or 29, as optimal dimension of training sample set will be $6 \times N$, and if number of hidden layer neurons is 30 or 31, as optimal dimension of training sample set will be $7 \times N$. In addition, the minimum values of the neural network mean square error 0.08, 0.15 and 0.24 are highlighted in the table 2.

Further experiments are reduced to comparing implemented neural network efficiency to efficiency of ready-made tools available on market. It should be noted that any instruments, able to recognize Ukrainian phonetics, are not currently available on market. However, an attempt to adapt tools, used for recognizing sounds in Russian, for recognition in Ukrainian was carried out. Ready tools usually recognize not individual sounds but whole words. These words are being formed by a sequence of sounds, therefore, the neural network proposed in the paper is able to cope with this task successfully. A spoken word is considered recognized incorrectly if its transcription differs from sequence of sounds got by neural network.

On the basis of comparing efficiency of attempts to adapt ready-made speech recognition tools to phonetic features in Ukrainian and tool, proposed in the paper, we can conclude advantages of the last one. So, in considered examples, relative improvement in quality of recognition is 12.3%. This is due to the initial orientation developed neural network to recognition speech in Ukrainian.

CONCLUSIONS

Got result is satisfactory, so neural network, proposed for recognizing sounds in Ukrainian, can be used to develop a full-fledged system for voice control on Android devices.

The scientific novelty of the obtained results is that the method for optimizing data of sound waves formed by pronunciation of sounds and their combinations in Ukrainian was firstly proposed. The proposed method is based on using only extreme values of sound waves characteristics and obtaining intermediate data by linear ap-

proximation. It reduced the training sample dimension and increased the data processing speed without losing quality of the result.

The practical significance of obtained results is that digitized sound waves, are generated during pronouncing separate sounds in Ukrainian, can be subjected to further intellectual analysis in order to search elements of certain device commands and their parameters. It is undoubtedly a very important element in building a full-fledged interface with voice control. However, considering this problem isn't within the scope of this study.

Prospects for further research are to train the proposed neural network to recognize whole words and collocations in Ukrainian and in other national languages.

REFERENCES

- Tohyama M., Tsunehiko K. Fundamentals of Acoustic Signal Processing. Boston, Academic Press, 1998, 321 p. DOI: 10.1121/1.429575
- Giannakopoulos T., Pirkakis A. Introduction to Audio Analysis: A MATLAB Approach. FL, Academic Press, 2014, 288 p.
- Lerch A. An introduction to audio content analysis. Applications in signal processing and music informatics. Hoboken, Wiley, 2012, 259 p. DOI: 10.1002/9781118393550
- Poulter C. Voice recognition software – Nuance Dragon naturally speaking, *Occupational Medicine*, 2020, Vol. 70, Issue. 1, pp. 75–76. <https://doi.org/10.1093/occmed/kqz128>
- Kumar A., Paek T., Lee B. Voice recognition software – Nuance Dragon naturally speaking/ *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA, 5–10 May 2012: proceedings*. Los Alamitos, IEEE, 2012, pp. 2277–2286.
- Natale S. To believe in Siri: A critical analysis of AI voice assistants, *Communicative Figurations Working Papers*, 2020, Vol. 32, pp. 130–146.
- Pham C., Yuto L., Yasuo T. A platform for integrating Alexa Voice Service into ECHONET-based smart homes, *Proceedings of International Conference on Consumer Electronics-Taiwan (ICCE-TW), Taichung, Taiwan, 19–21 May 2018: proceedings*. Taichung, IEEE, 2018, pp. 895–902. DOI: 10.1109/ICCE-China.2018.8448893
- Guzman A. L. Making AI Safe For Humans: A Conversation With Siri, *Socialbots and Their Friends: Digital Media and the Automation of Sociality/ edited by R. W. Gehl*. Routledge, 2017, Chapter 4, pp. 70–85.
- Khilari P., Bhoje V. P. A review on speech to text conversion methods, *International Journal of Advanced Research in Computer Engineering & Technology(IJARCET)*, 2015, Volume 4, Issue 7, pp. 3067–3072.
- Deepa V. J., Mustafa Alfateh, Sharan R. A Novel Model for Speech to Text Conversion, *International Refereed Journal of*

- Engineering and Science (IRJES)*, 2014, Volume 3, Issue 1, pp. 239–245.
11. Trivedi Ayushi, Pant Navya, Shah Pinal, Sonik Simran and Agrawal Supriya Speech to text and text to speech recognition systems-Areview, *IOSR Journal of Computer Engineering*, 2018, Volume 20, Issue 2, pp. 36–43. DOI: 10.9790/0661-2002013643.
 12. Saksamudre Suman K., Shrishrimal P. P., Deshmukh R. R. A Review on Different Approaches for Speech Recognition System, *International Journal of Computer Applications*, 2020, Volume 115, No. 22, pp. 385–396. DOI: 10.5120/20284-2839
 13. Chavan Rupali S., Sable Ganesh S. An Overview of Speech Recognition Using HMM, *International Journal of Computer Science and Mobile Computing*, 2013, Vol. 2, Issue 6, pp. 233–238.
 14. Furtună, T. F. Dynamic Programming Algorithms in Speech Recognition, *Revista Informatica Economică*, 2008, Vol. 2(46), pp. 94–98.
 15. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 6645–6649. DOI: 10.1109/ICASSP.2013.6638947
 16. Maas M. Lexicon-Free Conversational Speech Recognition with Neural Networks, *NAAC*, 2015, pp. 156–163. DOI: 10.3115/v1/N15-1038
 17. Ossama A.-H., Mohamed A. R., Jiang H. etc. Convolutional Neural Networks for Speech Recognition, *ACM transactions on audio, speech, and language processing*, 2014, Vol. 22, No. 10, pp. 1533–1545. DOI: 10.1109/TASLP.2014.2339736
 18. Lekshmi K. R., Sherly E. Automatic Speech Recognition using different Neural Network Architectures – A Survey, *International Journal of Computer Science and Information Technologies*, 2016, Vol. 7 (6), pp. 242–248.
 19. Aha D. W., Kibler D., Albert M. K. Instance-based learning algorithms, *Machine Learning*, 1991, No. 6, pp. 37–66. DOI: 10.1023/A:1022689900470

Received 14.07.2020.

Accepted 27.08.2020.

УДК 004.934.8:004.52

АРХІТЕКТУРА ТА АЛГОРИТМ НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ ГОЛОСОВИХ СИГНАЛІВ

Молчанова В. С. – канд. техн. наук, доцент кафедри інформатики ДВНЗ «Приазовський державний технічний університет», Маріуполь, Україна.

Міроненко Д. С. – канд. техн. наук, завідувач кафедри інформатики ДВНЗ «Приазовський державний технічний університет», Маріуполь, Україна.

АНОТАЦІЯ

Актуальність. Зазвичай взаємодія користувача з мобільним пристроєм, наприклад, телефоном або планшетом реалізується за допомогою торкань. Однак можливий цілий ряд ситуацій, коли здійснення такого способу людино-машинної взаємодії виявляється скрутним або навіть неможливим. Наприклад, при деяких захворюваннях опорно-рухового апарату можливе порушення моторики рухів, що в свою чергу призводить до неможливості повноцінної використовувати пристрій, помилок, втрати часу. У такій ситуації актуальним стає пошук альтернативних шляхів взаємодії користувача з системою. Розробка голосових інтерфейсів є одним з найбільш перспективних напрямків даної роботи.

Мета дослідження полягає в розробці методу оптимізації даних звукових хвиль і їх застосуванні при навчанні нейронної мережі для розпізнавання голосових сигналів, утворених вимовою звуків українською мовою.

Метод. Для реалізації проекту запропоновано у роботі системи, використовується підхід, заснований на аналізі аудіосигналу за формою утвореною їм звукової хвилі і спектрограми, а також застосуванні штучних нейронних мереж у процесі подальшої класифікації та виділення окремих, характерних для української мови, звуків. Нейронна мережа являє собою тришаровий перцептрон, структура якого повністю адаптована під особливості української фонетики. Враховується характер звукової хвилі, яка утворюється під час вимови того чи іншого звуку, а також кількість різноманітних звуків в українській фонетичі.

Результати. Проведено ряд експериментів, спрямованих на вибір оптимальної архітектури нейронної мережі і розмірність навчальної вибірки. В якості основного критерію при оцінці ефективності нейронної мережі використовувалося середньоквадратичне відхилення її помилки. В процесі тестування було визначено кілька варіантів комбінацій параметрів нейронної мережі, при яких досягалися найкращі результати. Порівняльний аналіз ефективності запропонованої в роботі нейронної мережі й існуючих на ринку інструментів розпізнавання голосу показав поліпшення відносних показників розпізнавання на 9,26%.

Висновки. Отримані в роботі результати досліджень і архітектура нейронної мережі можуть бути використані під час реалізації повноцінного голосового інтерфейсу для мобільних пристроїв під управлінням операційної системи Android. Незважаючи на те, що робота орієнтована на розпізнавання мовлення українською мовою, ідеї які використовуються для її реалізації можуть бути використані при транскрибації голосу на інших мовах.

КЛЮЧОВІ СЛОВА: голосовий інтерфейс, аудіосигнал, амплітуда сигналу, спектрограма, нейронна мережа, навчальна вибірка, середньоквадратичне відхилення.

УДК 004.934.8:004.52

АРХИТЕКТУРА И АЛГОРИТМ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ГОЛОСОВЫХ СИГНАЛОВ

Молчанова В. С. – канд. техн. наук, доцент кафедры информатики ГВУЗ «Приазовский государственный технический университет», Мариуполь, Украина.

Міроненко Д. С. – канд. техн. наук, заведующий кафедры информатики ГВУЗ «Приазовский государственный технический университет», Мариуполь, Украина.

АННОТАЦИЯ

Актуальность. Обычно взаимодействие пользователя с мобильным устройством, например, телефоном или планшетом реализуется посредством касаний. Однако возможен целый ряд ситуаций, когда осуществление такого способа человеко-

машинного взаємодіяння оказується затруднителним или даже невозможним. Например, при некоторых заболеваниях опорно-двигательного аппарата возможно нарушение моторики движений, что в свою очередь приводит к невозможности полноценно использовать устройство, ошибкам, потере времени. В сложившейся ситуации актуальным становится поиск альтернативных интерфейсов взаимодействия пользователя с системой. Разработка голосовых интерфейсов является одним из наиболее перспективных направлений данной работы.

Цель исследования состоит в разработке метода оптимизации данных звуковых волн и их применени при обучении нейронной сети для распознавания голосовых сигналов, образованных произношением звуков на украинском языке.

Метод. Для реализации проекта предложенной в работе системы, используется подход, основанный на анализе аудио-сигнала по форме образуемой им звуковой волны и спектрограммы, а также применении искусственных нейронных сетей в процессе последующей классификации и выделении отдельных, характерных для украинской речи, звуков. Нейронная сеть представляет собой трехслойный перцептрон, структура которого полностью адаптирована под особенности украинской фонетики. Учитывается характер звуковой волны, образуемой при произношении того или иного звука, а также количество разнообразных звуков в украинской фонетике.

Результаты Проведен ряд экспериментов, направлен на выбор оптимальной архитектуры нейронной сети и размерность обучающей выборки. В качестве основного критерия при оценке эффективности нейронной сети использовалось среднеквадратическое отклонение ее ошибки. В процессе тестирования были определены несколько вариантов комбинаций параметров нейронной сети, при которых достигались наилучшие результаты. Сравнительный анализ эффективности предложенной в работе нейронной сети и существующих на рынке инструментов распознавания показал улучшение относительных показателей распознавания на 9.26 %.

Выводы. Полученные в работе результаты исследований и архитектура нейронной сети могут быть использованы при реализации полноценного голосового интерфейса для мобильных устройств, работающих под управлением операционной системы Android. Несмотря на то, что работа ориентирована на распознавание речи на украинском языке, используемые при ее реализации идеи могут быть использованы при транскрипции речи на других языках.

КЛЮЧЕВЫЕ СЛОВА: голосовой интерфейс, аудиосигнал, амплитуда сигнала, спектрограмма, нейронная сеть, обучающая выборка, среднеквадратическое отклонение.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Tohyama M. Fundamentals of Acoustic Signal Processing / M. Tohyama K. Tsunehiko. – Boston : Academic Press, 1998. – 321 p. DOI: 10.1121/1.429575
2. Giannakopoulos T. Introduction to Audio Analysis: A MATLAB Approach / T. Giannakopoulos, A. Pikrakis. – FL: Academic Press, 2014. – 288 p.
3. Lerch A. An introduction to audio content analysis. Applications in signal processing and music informatics / A. Lerch. – Hoboken : Wiley, 2012. – 259 p. DOI: 10.1002/9781118393550
4. Poulter C. Voice recognition software – Nuance Dragon naturally speaking/ C. Poulter // Occupational Medicine. – 2020. – Vol. 70, Issue 1. – P. 75–76. <https://doi.org/10.1093/occmed/kqz128>
5. Voice recognition software – Nuance Dragon naturally speaking/ A. Kumar, T. Paek, B. Lee // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA, 5–10 May 2012 : proceedings. Los Alamitos: IEEE, 2012. – P. 2277–2286.
6. Natale S. To believe in Siri: A critical analysis of AI voice assistants/ S. Natale // Communicative Figurations Working Papers. – 2020. – Vol. 32. – P. 130–146.
7. Pham C. A platform for integrating Alexa Voice Service into ECHONET-based smart homes / C. Pham, L. Yuto, T. Yasuo // Proceedings of International Conference on Consumer Electronics-Taiwan (ICCE-TW), Taichung, Taiwan, 19–21 May 2018: proceedings. Taichung : IEEE, 2018. – P. 895–902. DOI: 10.1109/ICCE-China.2018.8448893
8. Guzman A. L. Making AI Safe For Humans: A Conversation With Siri / A. L. Guzman // Socialbots and Their Friends: Digital Media and the Automation of Sociality/ edited by R. W. Gehl. – Routledge, 2017. – Chapter 4. – P 70–85.
9. Khilari P. A review on speech to text conversion methods / P. Khilari, V. P. Bhope // International Journal of Advanced Research in Computer Engineering & Technology(IJARCET). – 2015. – Volume 4, Issue 7. – P. 3067–3072.
10. Deepa V. J. A Novel Model for Speech to Text Conversion / V. J. Deepa, Mustafa Alfateh, R. Sharan // International Reference Journal of Engineering and Science (IRJES). – 2014. – Volume 3, Issue 1. – P. 239–245.
11. Speech to text and text to speech recognition systems-Areview / [Ayushi Trivedi, Navya Pant, Pinal Shah et al] // IOSR Journal of Computer Engineering. – 2018. – Volume 20, Issue 2. – P. 36–43. DOI: 10.9790/0661-2002013643.
12. Saksamudre Suman K. A Review on Different Approaches for Speech Recognition System / Suman K. Saksamudre, P. P. Shrishrimal, R. R. Deshmukh // International Journal of Computer Applications. – 2020. – Volume 115, No. 22. – P. 385–396. DOI: 10.5120/20284-2839
13. Chavan Rupali S. An Overview of Speech Recognition Using HMM / Rupali S Chavan, Ganesh. S Sable. // International Journal of Computer Science and Mobile Computing. – 2013. – Vol. 2, Issue. 6. – P.233 – 238.
14. Furtună T. F. Dynamic Programming Algorithms in Speech Recognition/ T. F. Furtună // Revista Informatica Economică. – 2008. – Vol. 2(46). – P. 94–98.
15. Graves A. Speech recognition with deep recurrent neural networks / A. Graves, A. Mohamed, G. Hinton // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. – 2013. – P. 6645–6649. DOI: 10.1109/ICASSP.2013.6638947
16. Maas M. Lexicon-Free Conversational Speech Recognition with Neural Networks / M. Maas // NAAC. – 2015. – P. 156–163. DOI: 10.3115/v1/N15-1038
17. Ossama A.-H. Convolutional Neural Networksfor Speech Recognition / A.-H. Ossama, A. R. Mohamed, Jiang H etc. // ACM transactions on audio, speech, and language processing. – 2014. – Vol. 22, No. 10. – P. 1533–1545. DOI: 10.1109/TASLP.2014.2339736
18. Lekshmi K. R. Automatic Speech Recognition using different Neural Network Architectures – A Survey / K. R. Lekshmi, E. Sherly // International Journal of Computer Science and Information Technologies. – 2016. – Vol. 7 (6). – P. 242–248.
19. Aha D. W. Instance-based learning algorithms / D. W. Aha, D. Kibler, M. K. Albert // Machine Learning. – 1991. – № 6. – P. 37–66. DOI: 10.1023/A:1022689900470

THE GENERAL CONCEPT OF THE METHODS OF ALGORITHMIC CLASSIFICATION TREES

Povkhan I. F. – PhD, Associate Professor, Associate Professor at the Department of Software Systems, Uzhhorod National University, Uzhhorod, Ukraine.

ABSTRACT

Context. The general problem of constructing logical trees of recognition (classification) in the theory of artificial intelligence is considered in this paper. The object of this study is the concept of the classification tree (a logical and an algorithmic ones). The current methods and algorithms for constructing algorithmic classification trees are the subject of the study.

Objective. This work aims to create a simple and effective method for constructing tree-like recognition models on the basis of algorithmic classification trees for the training set of discrete information, which is characterized by the structure of the logical classification trees obtained on the basis of independent classification algorithms evaluated through the functional of calculating their overall efficiency.

Method. The general method of constructing algorithmic classification trees is proposed. It builds a tree-like structure (a classification model) for a given initial training data set. This structure consists of a set of autonomous algorithms of classification and recognition which have been evaluated at each step (stage) of constructing the model based on the given initial dataset. Namely, the method for constructing the algorithmic classification tree is proposed. The main idea of this method is to step by step approximate the initial dataset of arbitrary size and structure using a set of independent classification algorithms. This method, when forming the current vertex of the algorithmic tree (a node, a generalized feature) ensures the selection of the most effective (high-quality) autonomous classification algorithms from the initial dataset. In the process of constructing the resulting classification tree this approach can significantly reduce the size and complexity of the tree (the total number of branches, vertices and tiers of the structure) and improve the quality of its subsequent analysis (interpretability), the possibility of decomposition. The proposed method of constructing an algorithmic classification tree enables building different types of tree-like recognition models for a wide class of problems in the theory of artificial intelligence.

Results. The algorithmic classification tree method, developed and presented in this work, was implemented in the software and was studied and compared with the methods of logical classification trees (based on the selection of a set of elementary features) when solving the problem of recognizing real data of the geologic type.

Conclusions. The results of the conducted experiments described in this paper confirm the functional efficiency of the proposed mathematical software and show the possibility of its future use for solving a wide range of practical problems of recognition and classification. Further research prospects and approbation may consist in developing a limited method of the algorithmic classification tree, whose main points include the introduction of the criterion for stopping the procedure of constructing a tree model based on the depth of the structure, optimization of its software implementations, introduction of new types of algorithmic trees, and also the experimental research of this method while applying it for solving a wider range of practical problems.

KEYWORDS: algorithmic classification tree, pattern recognition, classification, classification algorithm, branching criterion.

ABBREVIATIONS

TS – training set;
CP – recognition system;
PR – pattern recognition;
GF – generalized feature;
RF – recognition function;
LCT – logical classification tree;
ACT – algorithmic classification tree;
GFT – generalized feature tree;
BFS – branch feature selection.

NOMENCLATURE

a_i is a fixed independent classification and recognition algorithms in the scheme of the algorithmic classification tree;

G is a some initial set of signals (discrete objects);

R is a partitioning into classes (patterns) H_i specified in the initial data set G ;

f_R is a recognition function (RF) defined in the initial data set G ;

x_i is a discrete objects (signals) of the initial TS;

H_i is a set of patterns (classes) specified in the initial TS;

$(x_i, f_R(x_i))$ is a training pairs of the initial TS;

m is a total number of training pairs (objects of the known classification) of the initial TS;

M is a total number of independent classification algorithms a_i in the set;

k is a total number of classes (patterns) of the set of signals G ;

n is a total number of features (attributes) of a problem (feature space dimension);

l is a value of class membership of discrete object x ;

G_{a_1, \dots, a_i} is a subset of the initial G which is approximated using the set of classification algorithms a_i ;

S_{a_1, \dots, a_n} is a number of occurrences in the TS of training pairs which meet the condition $x_i \in G_{a_1, \dots, a_n}$;

S_{a_1, \dots, a_n}^l is a number of all training pairs of the initial TS which are correctly classified (when the basic membership requirement is fulfilled $x_s \in G_{a_1, \dots, a_i}, f_R(x_s) = l$;

δ_{a_1, \dots, a_i} is a frequency of occurrence of the members of the sequence of discrete objects from the initial TS in G_{a_1, \dots, a_i} data set;

ψ_{a_1, \dots, a_i}^j is a frequency of membership of object x in pattern (class) H_j provided that $x \in G_{a_1, \dots, a_i}$;

ρ_{a_1, \dots, a_i} is a maximum frequency out of frequencies ψ_{a_1, \dots, a_i}^j regarding all the classes of the current problem;

F_S is a value which characterizes the quality of approximation of the initial TS using the set of classification algorithms a_i ;

F is a function of recognizing the structure of the classification tree;

m_F is a number of all training pairs $(x_j, f_R(x_j))$ of the initial TS which are classified correctly;

τ_F is a value which characterizes the efficiency of RF for the initial TS regarding some set of classification algorithms a_1, \dots, a_i ;

f_j is a some generalized feature constructed at j -step of generating the algorithmic classification tree model;

En_{tr} is a classification errors, rejections (of the first type) in the initial training data set;

Et_{tr} is a classification errors, rejections (of the first type) in the training data set;

T_{Uz} is a characterizes the total time (hardware time) which is necessary for building the current generalized feature f_j ;

E_{Uz} is a information capacity (structural complexity) of the constructed GF f_j within the current step of generating the algorithmic classification tree model;

S_{Uz} is a total number of discrete objects x_i of the initial TS which are generalized (described) by the specified generalized feature f_j ;

$P_{pt}(TS)$ is a power (volume) of the initial TS (or its fixed part for the current step of the scheme of constructing the algorithmic classification tree);

Er_{All} is a total number of errors of the algorithmic classification tree model in the data of the initial test set as well as of the training set;

M_{All} is a total power (volume) of the data of the training set as well as of the test set;

Fr_{All} is a characterizes the number of vertices of the obtained algorithmic classification tree model with the

resulting value f_R (RF, i. e. the leaves of the classification tree);

V_{All} is a represents the total number of all types of vertices in the structure of the algorithmic classification tree model;

O_{Uz} is a total number of generalized features used in the classification tree model;

P_{All} is a total number of transitions between the vertices in the structure of the constructed classification tree model;

N_{All} is a total number of different classification algorithms that are used in the classification tree model;

I_{Main} is a indicator of generalizing data of the initial TS using the classification tree;

Q_{Main} is a integral indicator of the quality of the algorithmic classification tree model

INTRODUCTION

Today, the rapid scientific and technological advancements urge an engineer to solve the fundamentally important problem that often arises when working with large amounts of data. This is the problem of efficient automatic construction of systems for processing large amounts of information, decision-making systems, data analysis systems. It is clear that the solution of this problem allows one to completely transfer the hard work related to designing a complex system to a computer and release an engineer's creative potential to solve other, more important and relevant problems. Moreover, overcoming this problem within the theory of artificial intelligence along with automating the algorithm and software construction of specific recognition systems in the form of LCT / ACT models is the key to their high efficiency for every real problem, and consequently will ensure the rapid development of various fields of science and technology [1].

Information technologies based on mathematical models of pattern recognition in the form of tree models are widely used in systems of processing and analyzing arrays of information. Apparently, this is due to the fact that this approach eliminates a set of shortcomings typical of classical methods and achieves a fundamentally new result efficiently and rationally using the power of computer information systems [2].

Today we know almost four thousand recognition methods and algorithms which are based on various approaches and concepts, but they all have certain limitations in their use – accuracy, speed of operation, memory. It should be underlined that each of the existing classification algorithms is limited to certain specifics of application problems (universality constraints), and this is certainly the weakest point of not only these algorithms, but also information systems in general, which are based on the respective concepts [3, 4].

Decision trees, namely the structures of algorithmic classification trees are **the object of the study**.

Thus, the representation of training sets (arrays of discrete information) containing large amount of data in the form of structures of logical (algorithmic) trees has its significant advantages in terms of economic description and analysis of data, efficient mechanisms (procedures) for working with them [4, 5]. The coverage of the TS with the set of elementary features in the case with the LCT or the coverage of the TS with the fixed set of autonomous recognition and classification algorithms in the case with the ACT, generates a fixed tree-like data structure (a tree model) which provides even compression and conversion of the initial training dataset. This approach enables the significant optimization and saving of hardware resources of the information system [6–9].

It is known that the field of application of the concept of decision trees (LCT / ACT) is currently extremely extensive; yet many tasks and problems which are solved with the help of this instrument can be reduced to the following three basic segments – problems of describing data structures, recognition and classification problems, regression problems [10]. The vast majority of modern schemes of methods for constructing classification trees are known in the literature as ‘divide and conquer’. It should be noted only that when this scheme is applied, the classification tree will be constructed in the direction from top to bottom [11].

An arbitrary structure of the classification tree (LCT / ACT) is presented in the form of branches and nodes, and on the branches of the tree there are some labels (attributes, values) on which the target function depends (in the case of the ACT – independent classification algorithms, sets of GFs), and the nodes (vertices) contain the RF values (the values of class membership) or the extended attributes of transitions.

The central issues of the concept of classification trees remain the ones linked to the choice of the branching criterion (constructing or selecting vertices), the branching stopping criterion (constructing the classification tree structure), and the criterion for rejecting branches of the logical tree (subtrees). This gives rise to the fundamental issue of the theory of classification trees which consists in the possible construction of all variants of logical trees that correspond to the initial TS and the selection of the minimal classification tree according to the depth, structural complexity (the number of tiers) [12–15].

The methods and algorithms for constructing algorithmic classification trees (decision trees) are **the subject of research**.

An important point about the existing methods of processing training sets (discrete arrays) in the recognition problems when classification rules are built is that they do not allow regulating their complexity (the parametric complexity of GFs) in the process of constructing the model [16–19]. This shortcoming is not found in the methods of constructing recognition systems which are based on the concept of algorithmic classification trees (decision trees). The peculiarity of the algorithmic tree method is the possibility of complex use for solving each specific problem of constructing the

scheme of recognition of many known recognition algorithms (methods). The concept of the ACT is based on the single methodology – the optimal approximation of the TS using the set of generalized features (autonomous algorithms), which are included in some scheme (operator), constructed in the training process [20].

The objective of the paper is to develop a simple, efficient and universal method of constructing models of classification (recognition) based on the concept of the ACT for discrete arrays, where the obtained schemes of classification systems are characterized by a tree-like structure and the presence of autonomous classification algorithms (the sets of GFs) as their structural elements.

1 PROBLEM STATEMENT

Suppose there is some initial TS as the sequence of training pairs which are the following:

$$(x_1, f_R(x_1)), \dots, (x_m, f_R(x_m)). \quad (1)$$

Here the objects $x_i \in G$ (G is some set), and RF $f_R(x_i) \in \{0, 1, 2, \dots, k-1\}$, ($i = 1, 2, \dots, m$). In addition to the initial TS, a test set (a set of objects of the known class membership) is also specified, as some part of the initial TS.

For RF $f_R(x_i) = l$, ($0 \leq l \leq k-1$) means that $x_i \in H_l$, $H_l \subset G$, and here f_R is some finite significant function which specifies the initial partitioning R of G set which consists of subsets (a set of patterns, classes) $H_0, H_1, H_2, \dots, H_{k-1}$.

Hence, according to the specifications, a TS is a set (fixed sequence) of some sets (discrete objects), and each set is a set of values of some features (attributes) as well as values of some functions (RFs) typical of this set. Then the set of values of the features is a certain image (a discrete object), and the value of the function (RF) refers this image to the corresponding pattern [21–23].

Thus, the paper deals with the problem of constructing the ACT model with p parameters, whose L structure is optimal $F(L(p, x_i), f_R(x_i)) \rightarrow opt$ in relation to the initial training dataset.

2 REVIEW OF THE LITERATURE

This research continues a series of works devoted to the problematics of tree-like recognition schemes (classification models) of discrete objects [4–7]. They highlight the issues of constructing, using and optimizing the structures of classification trees. As mentioned in [5], the resulting rule of classification (scheme), which is built with the help of an arbitrary method or algorithm of branched selection of features, has a tree-like logical structure, and the logical tree consists of vertices (features, attributes), which are grouped by tiers and which are obtained at a certain step (stage) of constructing the recognition tree [24]. An important task that arises from [20] is the one linked to synthesizing recognition

trees which will actually be represented by a tree (a graph) of algorithms (ACT methods). In contrast to the existing methods, the main feature of tree-like recognition systems is that the importance of individual features (groups of features or algorithms) is determined in reference to the function that specifies the partition of objects into classes [23]. Thus, work [15] is devoted to the principal issues concerning the generation of decision trees in the case of uninformative features, estimating the quality of the constructed models. The ability of classification tree structures to perform one-dimensional branching (the selection of features, attributes) for analyzing the impact (importance, quality) of individual variables (vertices) makes it possible to work with variables of different types in the form of predicates, generalized features, in the case of ACTs – with the respective autonomous classification and recognition algorithms. This concept of classification trees is actively used in data mining where the ultimate goal is to synthesize the model (fixed scheme), which predicts the value of the target variable based on the set of the initial data (training datasets) at the input of the system [26].

Nowadays a number of algorithms implementing the concept of decision trees (classification trees) are applied. However, their two representatives (C4.5 and CART) are widely spread and used; and also mentioned above the algorithm of the logical tree C4.5 as the node (vertex) selection criterion employs the so-called theoretical information criterion whereas CART algorithm is based on calculating Gini index which takes into account the relative distances (within the metric) between class distributions [27–30].

Since the main idea of methods and algorithms of the branched selection of features (vertices of algorithms) of ACTs can be defined as the optimal approximation of some initial TS using a set of ranked classification algorithms (features, attributes of the object in the case of LCTs), the central problem which arises is related to choosing an efficient branching criterion (the selection of vertices, attributes, features of discrete objects for LCTs and algorithms for ACTs). These fundamental problems are studied in [15, 21, 31–34]; they include quality evaluation of individual discrete features, their sets and fixed junctions; all of these enable introducing an efficient mechanism for the implementation of branching.

It should be noted that the structure of models of classification trees is characterized by compactness on the one hand and uneven occupancy (discharge) of the tiers on the other hand in comparison with regular trees [4]. It has to be stressed that the important issues, which still remain open, encompass the issue of the convergence of the process of constructing classification trees in accordance with the methods of branched feature selection and the issue of selecting the criterion for stopping the process of synthesizing a logical tree [19].

An important point is that there is no contradiction between the concept of classification trees and the possibility to use as the features (vertices of the structure) of the classification tree not only the individual attributes

(features) of objects of their junctions (the idea of the generalized feature which was studied in work [25]) and the sets; but if we go further and do not consider the attributes of objects (features) as branches – but select individual independent recognition algorithms (estimated based on the training dataset), a new structure – the ACT will be obtained at the output [20]. It is the ACT structures that this paper focuses on.

3 MATERIALS AND METHODS

By analogy with the methods of approximation of the TS using the set of estimated elementary features [3, 9, 15] – we present the main idea of the methods of algorithmic classification trees, which in turn lies in approximating the initial training dataset arrays with the help of the set of autonomous different-type algorithms of classification.

So, suppose there is the initial TS of general type (1) as a sequence of training pairs of the known classification (m power), and some system (set) of independent and autonomous recognition (classification) algorithms for the initial TS $a_1(x), a_2(x), \dots, a_M(x)$. Then it is necessary to introduce the following sets which represent the partition of the training dataset using the respective classification algorithms a_i :

$$G_{a_1, \dots, a_i} = \{x \in G / a_i(x) = 1\}, (i = 1, \dots, M). \quad (2)$$

It should be admitted that here in order to simplify the explanation each of the autonomous classification algorithms $a_i(x)$ generates at the output the values only within a binary set $\{0, 1\}$, id est here $a_i(x) = 1$ in case of the successful classification of object x and $a_i(x) = 0$ in the opposite case.

Let us note that the system of G_{a_1, \dots, a_i} sets actually represents the complete stepwise partitioning of G set (with the increasing i value – that is the involved classification algorithms), which is instantiated with the independent algorithms a_1, a_2, \dots, a_M . It must be stressed that depending on the initial selection of the set of classification algorithms a_1, a_2, \dots, a_M some of the sets G_{a_1, \dots, a_i} can be empty (the case of unsuitability of one specific or several algorithms for approximating the current TS).

At the next stage, by S_{a_1, \dots, a_M} value we mark the number of occurrences in the initial TS of those training pairs $(x_s, f_R(x_s))$, ($1 \leq s \leq m$) which meet the basic requirement of membership $x_s \in G_{a_1, \dots, a_i}$.

Respectively, by the value of S_{a_1, \dots, a_M}^j , ($j = 0, 1, \dots, k - 1$) we mark the number of occurrences in the TS of those pairs $(x_s, f_R(x_s))$, ($s = 1, 2, \dots, m$) which meet the requirements $x_i \in G_{a_1, \dots, a_n}$ and $f_R(x_s) = j$.

Therefore, taking into account the above-mentioned and by analogy with the methods of selecting elementary feature sets, we can introduce the following values, which should be considered as a certain criterion of branching in the ACT structure:

$$\delta_{a_1, \dots, a_i} = \frac{s_{a_1, \dots, a_i}}{m}; \quad (3a)$$

$$\psi_{a_1, \dots, a_i}^j = \frac{s_{a_1, \dots, a_i}^j}{s_{a_1, \dots, a_i}}; \quad (3b)$$

$$\rho_{a_1, \dots, a_i} = \max_j \psi_{a_1, \dots, a_i}^j. \quad (3c)$$

It should also be noted that if object $x_s \notin G_{a_1, \dots, a_i}$ for all $(s = 1, \dots, m)$, then it is clear that $\delta_{a_1, \dots, a_i} = 0$ and $\psi_{a_1, \dots, a_i}^j = 0$ when $(j = 0, 1, \dots, k-1)$.

Hence, here δ_{a_1, \dots, a_i} value characterizes the frequency of occurrences of the members of the sequence x_1, x_2, \dots, x_m (discrete objects) in G_{a_1, \dots, a_i} set, and correspondingly ψ_{a_1, \dots, a_i}^j value characterized the frequency of membership of the object of some x pattern (class) H_j if $x \in G_{a_1, \dots, a_i}$. It must be underscored here that the requirement $x \in G_{a_1, \dots, a_i}$ is equivalent to the requirement that in the sequence of algorithms a_1, \dots, a_i there will be identified such an algorithm a_y in which $a_y(x) = 1$. Then the value δ_{a_1, \dots, a_i} characterizes information efficiency of recognizing the membership of some object x in one of classes H_0, H_1, \dots, H_{k-1} certainly, provided that $x \in G_{a_1, \dots, a_i}$.

At the following stage again there arises the principal issue connected with the membership of object x in classes H_0, H_1, \dots, H_{k-1} (the issue of the formation of classification rules). It is apparent that object x should also be referred to the class H_j for which the simple ratio is fulfilled:

$$\rho_{a_1, \dots, a_i} = \psi_{a_1, \dots, a_i}^j. \quad (4)$$

We must stress that here $\{0 \leq j \leq k-1\}$, and the ratio (4) represents some classification rule, besides it is clear – the larger ρ_{a_1, \dots, a_i} value, the higher its efficiency.

Since the initial TS is the only information which represents the partitioning of patterns H_0, H_1, \dots, H_{k-1} , and under H_j class we understand a set of all training pairs $(x_s, f_R(x_s))$ of the TS which meet the ratio $f_R(x_i) = j$, id est the requirement of membership. Similarly, the average efficiency of recognizing the set of patterns H_0, H_1, \dots, H_{k-1} which are specified by the training

dataset using recognition algorithms a_1, a_2, \dots, a_M , is estimated by the following value:

$$F_S(a_1, \dots, a_M) = \sum_{a_1, \dots, a_i} \delta_{a_1, \dots, a_i} \cdot \rho_{a_1, \dots, a_i}. \quad (5)$$

Hence, in this case $F_S(a_1, \dots, a_M)$ value can be calculated by means of the estimation of approximation of the initial TS using the set of independent classification algorithms a_1, a_2, \dots, a_M .

In view of the idea of the algorithmic classification tree which has been introduced above $F_S(a_1, \dots, a_M)$ value can be obtained on the basis of the following speculations – let us say that recognition function F is some image which puts some element of set $\{0, 1, \dots, k-1\}$ (i. e., the corresponding number of the class) in line with each set a_1, \dots, a_i .

Thus, RF F is the function of type $F(a_1, \dots, a_i)$ where a_1, \dots, a_i take values from $\{0, 1\}$ set. According to RF $F(a_1, \dots, a_i)$ object $x, (x \in G)$ definitely refers to that class (pattern) $H_j, (0 \leq j \leq k-1)$ for which the following ratio is fulfilled:

$$F(a_1, \dots, a_i) = l, (0 \leq l \leq k-1). \quad (6)$$

Suppose there is the initial TS of type (1), then let us consider that RF $F(a_1, \dots, a_i)$ classifies set $(x_j, f_R(x_j)), (1 \leq j \leq m)$ of the data array (1) correctly if $F(a_1(x_j), \dots, a_i(x_j)) = f_R(x_j)$, in the opposite case RF $F(a_1, \dots, a_i)$ classifies $(x_j, f_R(x_j))$ training pair incorrectly.

Next, suppose m_F – is the number of all occurrences of training pairs $(x_j, f_R(x_j))$ in the initial TS which classify RF $F(a_1, \dots, a_i)$ correctly. Then we introduce the following value:

$$\tau_F = \frac{m_F}{m}. \quad (7)$$

It should be said that this value τ_F can be regarded as the total efficiency of RF $F(a_1, \dots, a_i)$ for the initial TS in reference to some set of classification algorithms a_1, \dots, a_i .

At the next stage of the research let us express this value τ_F through the previously proposed values δ_{a_1, \dots, a_i} and ψ_{a_1, \dots, a_i}^j . To do this let us count the quantity of those training pairs $(x_s, f_R(x_s))$ which are correctly classified by RF $F(a_1, \dots, a_i)$ and for which the ratio of membership $x_s \in G_{a_1, \dots, a_i}$ is fulfilled.

Suppose $F(a_1, \dots, a_i) = l$, then the number of all training pairs $(x_s, f_R(x_s))$ which are correctly classified (pro-

vided the basic requirement of membership is met $x_s \in G_{a_1, \dots, a_i}, f_R(x_s) = l$ is equal to S_{a_1, \dots, a_i}^l .

Then it becomes evident that m_F value, introduced above, is calculated according to the following formula:

$$m_F = \sum_{0 \leq a_1, \dots, a_i \leq 1} S_{a_1, \dots, a_i}^{F(a_1, \dots, a_i)}. \quad (8)$$

It should be noted here that in this formula (8) we can take into account only those training sets for which the ratio $S_{a_1, \dots, a_i} \neq 0$ is fulfilled. Thus, in view of the comment concerning (8) that has just been indicated we can introduce m_F value in the following way:

$$m_F = \sum_{0 \leq a_1, \dots, a_i \leq 1} S_{a_1, \dots, a_i} \cdot \psi_{a_1, \dots, a_i}^{F(a_1, \dots, a_i)}. \quad (9)$$

Let us admit that here $\psi_{a_1, \dots, a_i}^{F(a_1, \dots, a_i)}$ value was predefined in the expressions (4).

At the next stage the following is obtained from the formulae (9) and (8):

$$\tau_F = \frac{m_F}{m} = \sum_{0 \leq a_1, \dots, a_i \leq 1} \delta_{a_1, \dots, a_i} \cdot \psi_{a_1, \dots, a_i}^{F(a_1, \dots, a_i)}. \quad (10)$$

It should be emphasized here that since $\psi_{a_1, \dots, a_i}^{F(a_1, \dots, a_i)} \leq \rho_{a_1, \dots, a_i}$, then we shall have the following situation:

$$\tau_F \leq F_S(a_1, a_2, \dots, a_M). \quad (11)$$

Further, the recognition function (5) is denoted by F_0 , and this function is specified by the ratio $F_0(a_1, a_2, \dots, a_n) = l$ if:

$$\psi_{a_1, \dots, a_i}^l = \rho_{a_1, \dots, a_i} = \max_{0 \leq j \leq k-1} \psi_{a_1, \dots, a_i}^j. \quad (12)$$

Then based on the formulae (10) and (12) we immediately get the following:

$$\begin{aligned} \tau_{F_0} &= \frac{m_{F_0}}{m} = \sum_{0 \leq a_1, \dots, a_M \leq 1} \delta_{a_1, \dots, a_M} \cdot \psi_{a_1, \dots, a_i}^{F_0(a_1, \dots, a_i)} = \\ &= \sum_{0 \leq a_1, \dots, a_M \leq 1} \delta_{a_1, \dots, a_M} \cdot \rho_{a_1, \dots, a_M} = F(a_1, \dots, a_M). \end{aligned} \quad (13)$$

Hence, for all RFs $F(a_1, \dots, a_M)$ it follows from the formulae (13) and (11) that $\tau_F \leq \tau_{F_0}$.

It should be admitted that the result of functioning of each of the fixed (selected from the library of algorithms of some information system) autonomous classification and recognition algorithms a_i at the respective step of generating the ACT is one or several generalized features f_j (certain classification rules) which, in fact, describe (approximate) the determined part of the initial training set. Thus, the respective resulting generalized features for

the case of the known geometric recognition algorithms [2] are geometric objects that cover the TS in feature space of the n dimension problem.

It is clear that in real examples there may be the cases when the respective classification algorithm a_i cannot construct the generalized feature f_j due to the complex arrangement of H_k classes in the feature space of the problem or specific conceptual and implementation constraints of the classification algorithm itself. Then, by analogy with the LCT such a case is possible when the generalized features f_j , constructed with the help of a_i classification algorithm, approximate the initial TS incompletely or such a situation is supported by the algorithm scheme of generating the ACT (as an example, the initial restriction in the algorithm scheme of the classification tree – about generating no more than one generalized feature f_j at each stage of constructing the ACT model).

It must be underlined that the objects of the initial TS, which do not fall under the constructed scheme of approximating the set using the sequence of generalized features f_j (at the final stage of the procedure of synthesizing the ACT), are referred to as rejections (errors) of classification of the first type – En_{tr} and similarly for the test dataset the incorrectly classified discrete objects are also referred to as errors of the first type – Et_{tr} .

Therefore, given all the above, we can assume that the ACT structure (type I) will have the general structure of the type – (Fig. 1), where each tier of such a classification tree determines the stage of constructing the ACT by means of approximating some part of the TS making use of the current classification algorithm a_i and owing to this approach enables adjusting the final complexity (accuracy) of the obtained classification tree model.

It should be stressed that within each step of generating the ACT model – (Fig. 1) there are given the specific classification algorithm a_i and respective TS (or the subset of the initial TS), and the initial TS in full is provided only at the first step, further with the subsequent steps of constructing a classification tree the power of the training dataset arrays will decrease due to the set of the constructed GFs f_j which will cut off (describe) some part of the initial training dataset. It is also important to emphasize that depending on the structure of the ACT construction scheme and the peculiarities of the current algorithm a_i more than one GF f_j can be generated within each step.

At the next stage of the research for the ACT method we introduce two basic criteria for constructing the classification tree model – the criterion for stopping the procedure of branching K_{stop} (it actually adjusts the complexity and accuracy of the obtained ACT model) and

the criterion for selecting branches $W(a)$ (the choice of a classification algorithm at the current step) for the classification tree under construction.

Thus, based on the above, it is advisable to introduce K_{stop} – the criterion for stopping the branching process of the type (*boolean*) of the ACT construction procedure which consists in checking the power $P_{pt}(TS)$ of the training dataset of the following type:

$$K_{Stop} = \begin{cases} 0, & \text{if } P_{pt}(TS) = 0; \\ 1, & \text{if } P_{pt}(TS) > 0. \end{cases} \quad (14)$$

Let us admit that the procedure of constructing a classification tree continues unless $K_{Stop} = 1$, and the opposite situation when $K_{Stop} = 0$ signals about finishing the stage of synthesizing the ACT model and the necessity of transition to the step of the test check based on the test dataset and the estimation of the quality of the obtained classification tree model.

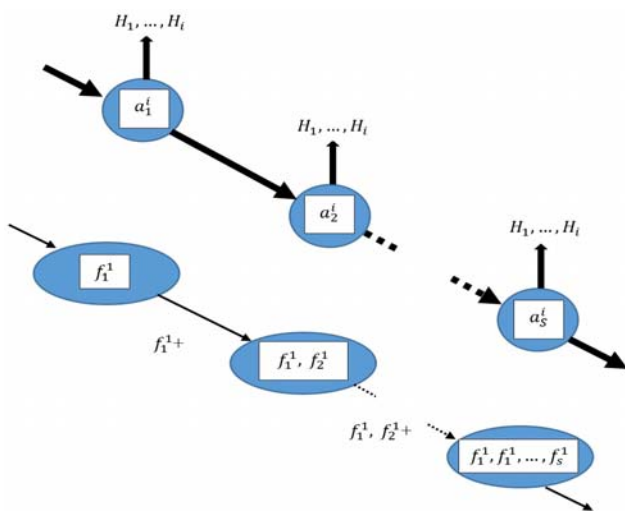


Figure 1 – The ACT structure of type I

It should be noted that while dealing with the ACT methods there arises the principal issue of selecting the branching criterion (the estimation and selection of classification algorithm a_i for the current step) in the structure of the classification tree model which is constructed. It is clear that similarly to the method of approximating the TS using the set of the ranked elementary features by the branching criterion we can suggest the initial estimation of efficiency of the set of algorithms (a_1, \dots, a_i) of the following type:

$$W(a_i) = \frac{1}{P_{pt}(TS)} \cdot \frac{\sum_{j=1}^k (T_{Uz} + S_{Uz} + \frac{E_{Uz}}{S_{Uz}})}{k} \quad (15)$$

Let us highlight that in the formula (15) summation is carried out based on all classes which are provided by the data array of the initial TS (though there can be restrictions on summation which are associated with the

structure (parameters) of the algorithm of constructing a classification tree).

An important point in the scheme of building the ACT model (Fig. 1) is that at each step of the tree algorithm actually its fixed (one or more – depending on the structure of the ACT algorithm itself) GF f_j is constructed, and along with this their total number increases with each coming step of the classification tree algorithm, namely the ACT with the set of classification algorithms (a_1, \dots, a_m) generates (replicates) a tree-like structure – a tree of generalized features with the appropriate set of GFs (f_1, f_2, \dots, f_z) .

Therefore, in view of all the above, there can be offered one of the possible algorithmic schemes for constructing the ACT (type I).

Stage 1. At the first stage of constructing the ACT the set of autonomous classification and recognition algorithms (a_1, a_2, \dots, a_M) is fixed in the library of algorithms of the information system (it is selected in an interactive mode randomly or after the procedure of appropriate estimation of the quality (efficiency) based on the initial training dataset). Let us note that both are selected – classification algorithms as well as their quantity (M value) depending on the aspects of the application problem.

Stage 2. At the second stage of constructing the ACT the selected set of classification algorithms (a_1, a_2, \dots, a_M) is estimated and ranked on the basis of the functional (15) according to the training dataset in the set, respectively, with their efficiency. It should be stressed that here similarly to the LCT two options are possible – depending on the algorithmic scheme of constructing a classification tree:

a) The option when the estimation of efficiency and the ranking of the set of classification algorithms (a_1, a_2, \dots, a_M) are done only once within this stage, and then at each step of constructing the ACT the following algorithm a_i of the initial sequence is fixed to approximate the data. This approach significantly saves the hardware resources of the information system, but negatively affects the complexity of the obtained classification tree model.

b) The option when the estimation of the efficiency and the ranking of the set of classification algorithms (a_1, a_2, \dots, a_M) are done at each step of constructing the ACT according to the appropriate data of the subset (parts) of the initial TS in order to estimate and identify the highest quality (efficient) classification algorithm for this part of the TS (the step of generating the ACT). This approach enables completing the approximation of the TS in a fewer steps and obtaining a more economical structure of the ACT compared to option (a), however, it requires much more hardware resources of the information system for the second stage of the scheme of constructing the ACT and requires considerable attention and the introduction of a set of restrictions on the initial

selection of the set of classification algorithms (a_1, a_2, \dots, a_M) .

Stage 3. At the third stage of constructing the ACT scheme the initial vertex of the ACT – the classification algorithm a_i of the highest efficiency out of the sorted set (a_1, a_2, \dots, a_M) , is fixed, and at the input the initial TS in the form of the sequence of training pairs is provided. Algorithm a_i ensures the generating of one or several GFs f_i of the first tier (the number of GFs generated at each step are determined by the parameters of the algorithm of constructing the ACT) which approximate (classify) some part of the TS.

Stage 4. At the fourth stage the next most efficient classification algorithm a_i of the ranked sequence (a_1, a_2, \dots, a_M) is chosen as the vertex of the second tier and the procedure of constructing the GFs of the third stage is repeated with the only difference that at the input there is provided already the limited TS without training pairs which are approximated by the GF, the vertex of the first tier, etcetera. Thus, further the procedure of constructing the ACT will come down to repeating this stage for the next most efficient a_i algorithm sequence (a_1, a_2, \dots, a_M) , the constant clipping of the parts of the TS and checking the branch stop criterion (the empty TS), which actually signals the completion of the procedure of constructing the ACT model and obtaining at the output a tree of classification algorithms a_i as well as a tree of generalized features f_i .

It should be mentioned that there are other implementation options of the scheme for constructing the ACT of the first type, which differ from the proposed scheme by variations in the number of GFs that are built at each step, the criteria and sequence of stages for assessing the quality of classification algorithms, the possibility of using a limited number of algorithms (even one), the possibility of approximating each of the classes of the TS using the set of its selected algorithms, the possibility of varying the criterion for stopping branching in the ACT.

Finally, it should be emphasized that the main peculiarity of such a scheme for constructing the ACT is the possibility to adjust the accuracy of the classification tree model, which is constructed during the basic procedure of constructing the ACT, along with this the important point is the principal possibility of constructing the ACT model with the predetermined accuracy with respect to the data array of the initial TS. This possibility is achieved by limiting the number of steps of the ACT generation procedure, the system of restrictions on information capacity, the number and parameters of generalization (the area of the approximated TS) of the set of generalized features which are constructed at the appropriate stages of constructing the resulting classification tree.

4 EXPERIMENTS

It has to be noted that the scheme for building the ACT suggested in this work enables adjusting the complexity of the classification tree model under construction, building models with the predetermined accuracy, and the classification tree structure consists of different-type autonomous classification algorithms as constructing modules (components). Moreover, the task of selecting the classification tree model from the set of constructed ACTs for a specific problem is determined by a set of parameters that are of decisive importance in relation to the current application problem (the training dataset). It is apparent that in order to compare and select a specific ACT model from the fixed set, it is necessary to determine their most important parameters (feature space dimension, the number of vertices, transitions, algorithms, etc.) and to identify their error regarding the input array.

It is fundamentally important to consider the criteria of quality of the obtained classification tree models, which depend on the model error, the power of the initial data array of the TS, the size of the test set (the number of training pairs and the dimension of the attribute space of the problem), the number of model parameters, etcetera [15, 27].

$$Q_{Main} = \frac{Fr_{All}}{V_{All} \cdot O_{Uz} \cdot P_{All} \cdot N_{All}} \cdot e^{-\frac{Er_{All}}{M_{All}}} \quad (16)$$

Let us underline that this integral indicator of the quality of the ACT model will take values between zero and one. The lower it is, the worse the quality of the constructed classification tree is, and the higher the indicator, the better model is obtained.

An important indicator that characterizes the basic properties of the obtained ACT models is the general indicator of the generalization of the data of the initial TS using the classification tree; this indicator is calculated as follows:

$$I_{Main} = \frac{m \cdot O_{Uz}}{V_{All} + N_{All} + 2P_{All}} \quad (17)$$

The proposed evaluation of the quality of the classification tree model (ACT) reflects its basic parameters (characteristics) of classification trees and can be applied as the criterion for optimality in the procedure of evaluating an arbitrary tree-like recognition scheme, for example in the case of methods of constructing and selecting random LCTs from work [24] (taking into account their respective structural parameters). It must be noted that the fundamental point which still remains is linked to reducing the complexity of the classification tree structure, the parameters of consuming memory λ and CPU time τ . It is also necessary to maximize the parameter I_{Main} (the indicator of generalization of the ACT model) that enables obtaining the most optimal structure of the classification tree and actually provides the maximum data compression of the initial TS (to represent the array of the initial data by the logical tree whose structural complexity is minimal) [12].

In Uzhhorod National University there has been developed “Orion III” software complex for generating autonomous recognition systems where the algorithmic library of the system has 11 recognition algorithms, among which there is the above-mentioned algorithmic implementation of constructing the ACT. Its task was to build an autonomous recognition system based on geological data (the problem of separating oil-bearing strata). 22 elementary features were used to identify the objects. The TS contains information about the objects of the two classes, and at the stage of the examination the constructed classification system (the ACT model) should ensure the effective recognition of objects of the unknown classification regarding these two classes. Note that before starting work, the training set was automatically checked for correctness (search and removal of the same objects of different membership – errors of the first kind)

5 RESULTS

Thus, the TS contains information on partition into two classes, and in the training data array there were predominantly training pairs of the first class (oil-bearing strata) in proportion (1.5 / 1), and the TS itself consisted of 1250 objects, and the efficiency of the constructed recognition system (the ACT model) was evaluated using the test set of 240 objects. The test set was a separate part of the training set (consisting of discrete objects of the known classification). The arrays of the training and test sets were obtained on the basis of geological exploration in Transcarpathian region during the period from 2001 to 2011.

Hence, a fragment of the main results of the above experiments (the comparative tests of methods for constructing LCT / ACT models using the data array of this application problem) is presented in (Table 1), and along with this the synthesized models of different-type classification trees provided the required level of accuracy specified in the task, the processing speed and costs of working memory of the information system, but showed different structural complexity of the constructed classification trees (models) and the set of generalized features in the case of the algorithmic classification tree model.

6 DISCUSSION

It is worth highlighting that the proposed estimates of the quality of the ACT model fix the most important characteristics of the classification trees and can be used as the criterion for optimality in the procedure of constructing the ACT and final selection from the set of ACT models. Notably, the method of the algorithmic tree operates only with the ready (constructed) generalized features, and it does not matter at all by means of what algorithm (rule, method) they were obtained, and each of the schemes constructed with the help of the algorithmic tree method is the general recognition system (the ATC model), which can be employed for practical work (processing large amounts of experimental data in the form of discrete sets). It should also be stressed that the resulting scheme is to some extent a new recognition algorithm (certainly, the one which has been synthesized from the known algorithms and methods). It is necessary to admit that the obtained ACT structure is characterized by the high degree of universality and relatively compact structure of the model itself, but it requires large hardware costs to store the generalized features of the initial assessment of the quality of classification algorithms according to the TS.

CONCLUSIONS

The problem of automation of constructing the models of algorithmic classification trees on the basis of approximating the TS using the set of independent classification algorithms has been solved in this paper.

The scientific novelty of the obtained results lies in the fact that for the first time there has been suggested a simple method of constructing the ACT based on evaluating and ranking the set of autonomous recognition algorithms for generating the classification tree structure (the ACT model). Moreover, within each step of branching the classification tree, a certain part of the TS (or its subset) is approximated. Admittedly, the functional (branching criterion) proposed in the work can be used not only for estimating the quality of individual classification algorithms, but also calculating the efficiency of the related sets of algorithms, which in the long run enables obtaining a more optimal structure of the synthesized ACT based on the initial training dataset.

Table 1 – The comparative table of the models / methods of classification trees.

No	The method of synthesizing the logical tree structure	Integral indicator of the model quality Q_{Main}	The total number of errors of the model in the training dataset and test dataset Er_{All}
1	The complete LCT method based on the selection of elementary features BFS-I	0.004691	2
2	The LCT method with one-time estimation of feature importance BFS-II	0.002172	3
3	The constrained BFS – LCT construction method	0.003219	4
4	The algorithmic tree method (type I)	0.005234	0
5	The algorithmic tree method (type II)	0.002941	0

The paper offers the set of general indicators (parameters) which allows the efficient presentation of the general characteristics of the ACT model, its possible use for selecting the most optimal ACT.

The applied value of the obtained results is that the proposed method of constructing ACT models was implemented in the library of algorithms of the universal software system “ORION III” to solve various practical problems of classification (recognition) of arrays of discrete objects. It must be underscored that the conducted practical tests confirm the functionality of the proposed ACT models and the developed software, which enables making a recommendation on the use of this approach and its software implementation for a wide range of application problems of classifying and recognizing discrete objects.

Further research is needed to develop methods of algorithmic classification trees (constrained methods) of optimizing software implementations of the proposed method for constructing the ACT as well as its practical approbation dealing with a number of real problems of classification and recognition.

REFERENCES

1. Srikant R., Agrawal R. Mining generalized association rules, *Future Generation Computer Systems*, 1997, Vol. 13, No. 2, pp. 161–180.
2. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Stanford, 2008, 768 p.
3. Quinlan J.R. Induction of Decision Trees, *Machine Learning*, 1986, No. 1, pp. 81–106.
4. Vasilenko Y. A., Vasilenko E. Y., Kuhayivsky A. I., Papp I. O. Construction and optimization of recognizing systems, *Scientific and technical journal “Information technologies and systems”*, 1999, No. 1, pp. 122–125.
5. Povhan I. Designing of recognition system of discrete objects, 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), 2016, Lviv, Ukraine. Lviv, 2016, pp. 226–231.
6. Mitchell T. Machine learning. New York, McGraw-Hill, 1997, 432 p.
7. Povhan I. General scheme for constructing the most complex logical tree of classification in pattern recognition discrete objects, *Collection of proceedings «Electronics and information technology»*, 2019, Vol. 11, pp. 73–80.
8. Breiman L. L., Friedman J. H., Olshen R. A., Stone C. J. Classification and regression trees. Boca Raton, Chapman and Hall/CRC, 1984, 368 p.
9. Vasilenko Y. A., Vashuk F. G., Povkhan I. F. Automating the construction of classification systems based on agent – schemes, *Mathematical modeling, optimization and information technologies : International Joint Conference MDIF-2012*. Kishineu, Moldova, 2012, Kishineu, 2012, pp. 444–446.
10. Vtoghoff P. E. Incremental Induction of Decision Trees, *Machine Learning*, 1989, № 4, pp. 161–186.
11. Amit Y., Geman D., Wilder K. Joint induction of shape features and tree classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, Vol. 19, No. 11, pp. 1300–1305.
12. Dietterich T. G., Kong E. B. Machine learning bias, statistical bias, and statistical variance of decision tree algorithms [Electronic resource]. Corvallis, Oregon State University, 1995, 14 p. Access mode : <http://www.cems.uwe.ac.uk/~irjohnso/coursenotes/uqc832/trbias.pdf>
13. Mingers J. An empirical comparison of pruning methods for decision tree induction, *Machine learning*, 1989, Vol. 4, No. 2, pp. 227–243.
14. Povhan I. Question of the optimality criterion of a regular logical tree based on the concept of similarity, *Collection of proceedings “Electronics and information technology”*, 2020, Vol. 13, pp. 12–16.
15. Subbotin S.A. Construction of decision trees for the case of low-information features, *Radio Electronics, Computer Science, Control*, 2019, № 1, pp. 121–130.
16. Lupei M., Mitsa A., Repariuk V., Sharkan V. Identification of authorship of Ukrainian-language texts of journalistic style using neural networks, *Eastern-European Journal of Enterprise Technologies*, 2020, Vol. 1 (2 (103)), pp. 30–36. DOI: <https://doi.org/10.15587/1729-4061.2020.195041>
17. Bodyanskiy Y., Vynokurova O., Setlak G. and Pliss I. Hybrid neuro-neo-fuzzy system and its adaptive learning algorithm, *Computer Sciences and Information Technologies (CSIT) : Xth Scien. and Tech. Conf. Lviv, 2015*. Lviv, 2015, P. 111–114.
18. Karimi K., Hamilton H. J. Generation and Interpretation of Temporal Decision Rules, *International Journal of Computer Information Systems and Industrial Management Applications*, 2011, Vol. 3, pp. 314–323.
19. Kotsiantis S. B. Supervised Machine Learning: A Review of Classification Techniques, *Informatica*, 2007, No. 31, pp. 249–268.
20. Povkhan I. F. Features of synthesis of generalized features in the construction of recognition systems using the logical tree method, *Information technologies and computer modeling ITKM-2019 : materials of the international scientific and practical conference, Ivano-Frankivsk, May 20–25, 2019*. Ivano-Frankivsk, 2019, pp. 169–174.
21. Vasilenko Y. A., Vashuk F. G., Povkhan I. F. The importance of discrete signs, *XX International Conference Promising ways and directions of improving the educational system, Uzhgorod, November 16–19, 2010*. Uzhgorod, 2010, Vol. 21, No. 1, pp. 217–222.
22. Deng H., Runger G., Tuv E. Bias of importance measures for multi-valued attributes and solutions, *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN)*. Espoo, Finland, Jun 14–Jun 17, 2011, Espoo, 2011, pp. 293–300.
23. Kamiński B., Jakubczyk M., Szufel P. A framework for sensitivity analysis of decision trees, *Central European Journal of Operations Research*, 2017, Vol. 26 (1), pp. 135–159.
24. Dietterich T. G. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization, *Machine learning*, 2000, Vol. 40, № 2, pp. 139–157.

25. Povhan I. Generation of elementary signs in the general scheme of the recognition system based on the logical tree, *Collection of proceedings «Electronics and information technology»*, 2019, Vol. 12, pp. 20–29.
26. Subbotin S., Oliinyk A. eds. : Szewczyk R., Kaliczynska M. The dimensionality reduction methods based on computational intelligence in problems of object classification and diagnosis, *Recent Advances in Systems, Control and Information Technology*. Cham, Springer, 2017, pp. 11–19. (Advances in Intelligent Systems and Computing, vol. 543).
27. Subbotin S. A. Methods and characteristics of localitypreserving transformations in the problems of computational intelligence, *Radio Electronics, Computer Science, Control*, 2014, No. 1, pp. 120–128.
28. Koskimaki H., Juutilainen I., Laurinen P., Roning J. Two-level clustering approach to training data instance selection: a case study for the steel industry, *Neural Networks : International Joint Conference (IJCNN-2008)*. Hong Kong, 1–8 June 2008 : proceedings. Los Alamitos, IEEE, 2008, pp. 3044–3049. DOI: 10.1109/ijcnn.2008.4634228
29. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition, *Optical Memory and Neural Networks (Information Optics)*, 2013, Vol. 22, № 2, pp. 97–103. DOI: 10.3103/s10660992x13020082
30. Subbotin S.A. Methods of sampling based on exhaustive and evolutionary search, *Automatic Control and Computer Sciences*, 2013, Vol. 47, № 3, pp. 113–121. DOI: 10.3103/s0146411613030073
31. De Mántaras R. L. A distance-based attribute selection measure for decision tree induction, *Machine learning*, 1991, Vol. 6, № 1, pp. 81–92.
32. Alpaydin E. Introduction to Machine Learning. London, The MIT Press. 2010, 400 p.
33. Painsky A., Rosset S. Cross-validated variable selection in tree-based methods improves predictive performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, Vol. 39, No. 11, pp. 2142–2153. DOI:10.1109/tpami.2016.2636831
34. Miyakawa M. Criteria for selecting a variable in the construction of efficient decision trees, *IEEE Transactions on Computers*, 1989, Vol. 38, No. 1, pp. 130–141.

Received 25.05.2020.

Accepted 30.08.2020.

УДК 001.891:65.011.56

ЗАГАЛЬНА КОНЦЕПЦІЯ МЕТОДІВ АЛГОРИТМІЧНИХ ДЕРЕВ КЛАСИФІКАЦІЇ

Повхан І. Ф. – канд. техн. наук, доцент, доцент кафедри програмного забезпечення систем ДВНЗ Ужгородський національний університет, м. Ужгород, Україна.

АНОТАЦІЯ

Актуальність. Розглянута загальна задача побудови логічних дерев розпізнавання (класифікації) в теорії штучного інтелекту. Об'єктом даного дослідження є концепція дерева класифікації (логічного та алгоритмічного). Предметом дослідження є актуальні методи та алгоритми побудови алгоритмічних дерев класифікації.

Мета. Метою даної роботи є створення простого та ефективного методу побудови деревоподібних моделей розпізнавання на основі алгоритмічних дерев класифікації для навчальних вибірок дискретної інформації, який характеризується структурою отриманих логічних дерев класифікації з незалежних алгоритмів класифікації оцінених на основі функціоналу розрахунку їх загальної ефективності.

Метод. Пропонується загальний метод побудови алгоритмічних дерев класифікації, який для заданої початкової навчальної вибірки буде деревоподібною структурою (модель класифікації), яка складається з набору автономних алгоритмів класифікації та розпізнавання оцінених на кожному кроці (етапі) побудови моделі за даною початковою вибіркою. Тобто пропонується метод побудови алгоритмічного дерева класифікації основна ідея якого полягає в по кроковій апроксимації початкової вибірки довільного об'єму та структури набором незалежних алгоритмів класифікації. Даний метод при формуванні поточної вершини алгоритмічного дерева (вузла, узагальненої ознаки) забезпечує виділення найбільш ефективних (якісних) автономних алгоритмів класифікації з початкового набору. Такий підхід при побудові результуючого дерева класифікації дозволяє значно скоротити розмір та складність дерева (загальну кількість гілок, вершин та ярусів структури) підвищити якість його наступного аналізу (інтерпретабельність), можливість декомпозиції. Запропонований метод побудови алгоритмічного дерева класифікації дозволяє будувати різноманітні деревоподібні моделі розпізнавання для широкого класу задач теорії штучного інтелекту.

Результати. Розроблений та представлений в даній роботі метод алгоритмічного дерева класифікації отримав програмну реалізацію та був досліджений та порівняний з методами логічних дерев класифікації (на основі селекції набору елементарних ознак) при розв'язку задач розпізнавання реальних даних геологічного типу.

Висновки. Проведені в даній роботі експерименти підтвердили працездатність запропонованого математичного забезпечення та показують можливість його перспективного використання для розв'язку широкого спектру практичних задач розпізнавання та класифікації. Перспективи подальших досліджень та апробацій можуть полягати в створенні обмеженого методу алгоритмічного дерева класифікації, який полягає в веденні критерію зупинки процедури побудови моделі дерева за глибиною структури, оптимізації його програмних реалізацій, введення нових типів алгоритмічних дерев а також експериментальних дослідженнях даного методу на більш широке коло практичних задач.

КЛЮЧОВІ СЛОВА: алгоритмічне дерево класифікації, розпізнавання образів, класифікація, алгоритм класифікації, критерій розгалуження.

ОБЩАЯ КОНЦЕПЦИЯ МЕТОДОВ АЛГОРИТМИЧЕСКИХ ДЕРЕВЬЕВ КЛАССИФИКАЦИИ

Повхан И. Ф. – канд. техн. наук, доцент, доцент кафедры программного обеспечения систем ГВУЗ Ужгородский национальный университет, г. Ужгород, Украина.

АННОТАЦИЯ

Актуальность. Рассмотрена общая задача построения логических деревьев распознавания (классификации) в теории искусственного интеллекта. Объектом данного исследования является концепция дерева классификации (логического и алгоритмического). Предметом исследования являются актуальные методы и алгоритмы построения алгоритмических деревьев классификации.

Цель. Целью данной работы является создание простого и эффективного метода построения древовидных моделей распознавания на основе алгоритмических деревьев классификации для учебных выборок дискретной информации, который характеризуется структурой полученных деревьев классификации из независимых алгоритмов классификации оцененных на основе функционала расчета их общей эффективности.

Метод. Предлагается общий метод построения алгоритмических деревьев классификации, который для заданной начальной учебной выборки строит древовидную структуру (модель классификации), которая состоит из набора автономных алгоритмов классификации, оцененных на каждом шагу (этапе) построения модели по данной начальной выборке. То есть предлагается метод построения алгоритмического дерева классификации основная идея которого заключается в по шаговой аппроксимации начальной выборки фиксированного объема и структуры набором независимых алгоритмов классификации. Данный метод при формировании текущей вершины дерева классификации (узла, обобщенного признака) обеспечивает выделение наиболее эффективных (качественных) автономных алгоритмов классификации из начального набора. Такой подход при построении результирующего дерева классификации позволяет значительно сократить размер и сложность дерева (общее количество ветвей, вершин и ярусов структуры) повысить качество его последующего анализа (интерпретируемость), возможность декомпозиции. Предложенный метод построения алгоритмического дерева классификации позволяет строить разнотипные древовидные модели распознавания для широкого класса задач теории искусственного интеллекта.

Результаты. Разработанный и представленный в данной работе метод алгоритмического дерева классификации получил программную реализацию и был исследован и сравнен с методами логических деревьев классификации (на основе селекции набора элементарных признаков) при решении задачи распознавания реальных данных геологического типа.

Выводы. Проведенные в данной работе эксперименты подтвердили работоспособность предложенного математического обеспечения и показывают возможность его перспективного использования для решения широкого спектра практических задач распознавания и классификации. Перспективы дальнейших исследований и апробаций могут заключаться в создании ограниченного алгоритмического метода дерева классификации, который заключается в ведении критерия останова процедуры построения модели дерева по глубине структуры, оптимизации его программных реализаций, введение новых типов алгоритмических деревьев, а также экспериментальных исследованиях данного метода в более широкой области практических задач.

КЛЮЧЕВЫЕ СЛОВА: алгоритмическое дерево классификации, распознавание образов, классификация, алгоритм классификации, критерий разветвления.

ЛИТЕРАТУРА / LITERATURA

1. Srikant R. Mining generalized association rules / R. Srikant, R. Agrawal // *Future Generation Computer Systems*. – 1997. – Vol. 13, № 2. – P. 161–180.
2. Hastie T. *The Elements of Statistical Learning* / T. Hastie, R. Tibshirani, J. Friedman. – Stanford, 2008. – 768 p.
3. Quinlan J. R. *Induction of Decision Trees* / J. R. Quinlan // *Machine Learning*. – 1986. – № 1. – P. 81–106.
4. Construction and optimization of recognizing systems / [Y. A. Vasilenko, E. Y. Vasilenko, A. I. Kuhayivsky, I. O. Papp] // *Scientific and technical journal "Information technologies and systems"*. – 1999. – № 1. – P. 122–125.
5. Povhan I. Designing of recognition system of discrete objects / I. Povhan // *2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP)*, Lviv, 2016, Ukraine. – Lviv, 2016. – P. 226–231.
6. Mitchell T. *Machine learning* / T. Mitchell. – New York : McGrawHill, 1997. – 432 p.
7. Povhan I. General scheme for constructing the most complex logical tree of classification in pattern recognition discrete objects / I. Povhan // *Collection of proceedings "Electronics and information technology"*. – 2019. – Vol. 11. – P. 73–80.
8. *Classification and regression trees* / L. L. Breiman, J. H. Friedman, R. A. Olshen et al.]. – Boca Raton : Chapman and Hall/CRC, 1984. – 368 p.
9. Vasilenko Y. A. Automating the construction of classification systems based on agent – schemes / Y. A. Vasilenko, F. G. Vashuk, I. F. Povkhan // *Mathematical modeling, optimization and information technologies : International Joint Conference MDIF-2012*, Kishineu, Moldova, 2012. – Kishineu, 2012. – P. 444–446.
10. Vtoghoff P. E. *Incremental Induction of Decision Trees* / P. E. Vtoghoff // *Machine Learning*. – 1989. – №4. – P. 161–186.
11. Amit Y. Joint induction of shape features and tree classifiers / Y. Amit, D. Geman, K. Wilder // *IEEE*

- Transactions on Pattern Analysis and Machine Intelligence. – 1997. – Vol. 19, № 11. – P. 1300–1305.
12. Dietterich T. G. Machine learning bias, statistical bias and statistical variance of decision tree algorithms [Electronic resource] / T. G. Dietterich, E. B. Kong. – Corvallis : Oregon State University, 1995. – 14 p. Access mode : <http://www.cems.uwe.ac.uk/~irjohnso/coursenotes/uqc832/trbias.pdf>
 13. Mingers J. An empirical comparison of pruning methods for decision tree induction / J. Mingers // *Machine learning.* – 1989. – Vol. 4, № 2. – P. 227–243.
 14. Povhan I. Question of the optimality criterion of a regular logical tree based on the concept of similarity / I. Povhan // *Collection of proceedings «Electronics and information technology».* – 2020. – Vol. 13. – P. 12–16.
 15. Subbotin S. A. Construction of decision trees for the case of low-information features / S. A. Subbotin // *Radio Electronics, Computer Science, Control.* – 2019. – № 1. – P. 121–130.
 16. Identification of authorship of Ukrainian-language texts of journalistic style using neural networks / [M. Lupei, A. Mitsa, V. Repariuk, V. Sharkan] // *Eastern-European Journal of Enterprise Technologies.* – 2020. – Vol. 1 (2 (103)). – P. 30–36. DOI: <https://doi.org/10.15587/1729-4061.2020.195041>
 17. Hybrid neuro-neo-fuzzy system and its adaptive learning algorithm / [Y. Bodyanskiy, O. Vynokurova, G. Setlak and I. Pliss] // *Computer Sciences and Information Technologies (CSIT) : Xth Scien. and Tech. Conf., Lviv, 2015.* – Lviv, 2015. – P. 111–114.
 18. Karimi K. Generation and Interpretation of Temporal Decision Rules / K. Karimi, H. J. Hamilton // *International Journal of Computer Information Systems and Industrial Management Applications.* – 2011. – Vol. 3. – P. 314–323.
 19. Kotsiantis S. B. Supervised Machine Learning: A Review of Classification Techniques / S. B. Kotsiantis // *Informatica.* – 2007. – №31. – P. 249–268.
 20. Povkhan I. F. Features of synthesis of generalized features in the construction of recognition systems using the logical tree method / I. F. Povkhan // *Information technologies and computer modeling ITKM-2019 : materials of the international scientific and practical conference, Ivano-Frankivsk, May 20–25, 2019.* – Ivano-Frankivsk, 2019. – P. 169–174.
 21. Vasilenko Y. A. The importance of discrete signs / Y. A. Vasilenko, F. G. Vashuk, I. F. Povkhan // *XX International Conference Promising ways and directions of improving the educational system, Uzhgorod, November 16–19, 2010.* – Uzhgorod, 2010. – Vol. 21, № 1. – P. 217–222.
 22. Deng H. Bias of importance measures for multi-valued attributes and solutions / H. Deng, G. Runger, E. Tuv // *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN), Espoo, Finland, Jun 14–Jun 17, 2011.* – Espoo, 2011. – P. 293–300.
 23. Kamiński B. A framework for sensitivity analysis of decision trees / B. Kamiński, M. Jakubczyk, P. Szufel // *Central European Journal of Operations Research.* – 2017. – Vol. 26 (1). – P. 135–159.
 24. Dietterich T. G. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization / T. G. Dietterich // *Machine learning.* – 2000. – Vol. 40, № 2. – P. 139–157.
 25. Povhan I. Generation of elementary signs in the general scheme of the recognition system based on the logical tree / I. Povhan // *Collection of proceedings «Electronics and information technology».* – 2019. – Vol. 12. – P. 20–29.
 26. Subbotin S. The dimensionality reduction methods based on computational intelligence in problems of object classification and diagnosis / S. Subbotin, A. Oliinyk // *Recent Advances in Systems, Control and Information Technology* / eds.: R. Szewczyk, M. Kaliczyńska. – Cham : Springer, 2017. – P. 11–19. – (Advances in Intelligent Systems and Computing, vol. 543).
 27. Subbotin S. A. Methods and characteristics of locality-preserving transformations in the problems of computational intelligence / S. A. Subbotin // *Radio Electronics, Computer Science, Control.* – 2014. – № 1. – P. 120–128.
 28. Two-level clustering approach to training data instance selection: a case study for the steel industry / [H. Koskimaki, I. Juutilainen, P. Laurinen, J. Roning] // *Neural Networks : International Joint Conference (IJCNN-2008), Hong Kong, 1–8 June 2008 : proceedings.* – Los Alamitos : IEEE, 2008. – P. 3044–3049. DOI: [10.1109/ijcnn.2008.4634228](https://doi.org/10.1109/ijcnn.2008.4634228)
 29. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition / S. Subbotin // *Optical Memory and Neural Networks (Information Optics).* – 2013. – Vol. 22, № 2. – P. 97–103. DOI: [10.3103/s1060992x13020082](https://doi.org/10.3103/s1060992x13020082)
 30. Subbotin S. A. Methods of sampling based on exhaustive and evolutionary search / S. A. Subbotin // *Automatic Control and Computer Sciences.* – 2013. – Vol. 47, № 3. – P. 113–121. DOI: [10.3103/s0146411613030073](https://doi.org/10.3103/s0146411613030073)
 31. De Mántaras R. L. A distance-based attribute selection measure for decision tree induction / De Mántaras R. L. // *Machine learning.* – 1991. – Vol. 6, № 1. – P. 81–92.
 32. Alpaydin E. *Introduction to Machine Learning* / E. Alpaydin. – London : The MIT Press, 2010. – 400 p.
 33. Painsky A. Cross-validated variable selection in tree-based methods improves predictive performance / A. Painsky, S. Rosset // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* – 2017. – Vol. 39, № 11. – P. 2142–2153. DOI: [10.1109/tpami.2016.2636831](https://doi.org/10.1109/tpami.2016.2636831)
 34. Miyakawa M. Criteria for selecting a variable in the construction of efficient decision trees / M. Miyakawa // *IEEE Transactions on Computers.* – 1989. – Vol. 38, № 1. – P. 130–141.

REFERENCE POINTS METHOD FOR HUMAN HEAD MOVEMENTS TRACKING

Rakova A. O. – Student of the Software Engineering Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Bilous N. V. – PhD, Associate Professor, Professor of the Software Engineering Department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

ABSTRACT

Context. The direction of the human face vector is an indicator of human attention. It has many applications in our daily lives, such as human-computer interaction, teleconferencing, virtual reality and 3D sound rendering. Moreover, determining the position of the head can be used to compare the exercises performed by a person with a certain standard, which brings us to investigation of ways to efficiently track moves. Depth-camera based systems, frequently used for these purposes, have significant drawbacks such as accuracy decreasing on the direct sunlight and necessity of additional equipment. The recognition from the two-dimensional image becomes more widespread and eliminates difficulties related to depth cameras which allows them to be used indoors and outdoors.

Objective. The purpose of this work is creation of the method that will allow us to track human head moves and record only significant vectors of head direction.

Methods. This paper suggests reference points method that decreases set of recorded vectors to minimal amount significant to describe head moves. It also investigates and compares existing methods for determining the vector of the face in terms of use in suggested approach.

Results. Suggested reference points method shows ability to highly decrease set of head direction vectors that describe the move. According to the results of the study, regression-based methods showed significantly better accuracy and independence from light and partial face closure so they were chosen to be used as methods to get head direction vector in reference points approach.

Conclusions. Research confirmed applicability of reference points method for human movements tracking and shown that methods of determining human head vector by two-dimensional image can compete in accuracy with RGBD-based methods. Thus combined with suggested approach these methods expose less restrictions in use than RGBD-based ones.

KEYWORDS: Face orientation vector, head moves, recognition, deep learning.

ABBREVIATIONS

CNN is a convolutional neural network;
FPS is a Frames per Second;
MAE is a Mean absolute error;
PAM is a parameterized appearance models;
RGBD is RGB image with depth information.

NOMENCLATURE

\bar{a}_i is an angle vector on the i -th frame;
 \bar{a}'_i is a reference point vector;
 (c_x, c_y) is a focal center;
 e_i is an absolute error;
 f_x and f_y are focal lengths in the x and y directions;
 k is a flexibility degree;
 n is an amount of observations in experiment;
 r_i is an element of the rotation vector;
 t_i is an element of camera translation vector;
 U, V, W are positions of the object in three-dimensional space;
 x, y are coordinates of the points in the image;
 x_i is a real value;
 y_i is a predicted value.

INTRODUCTION

Human face direction as a part of entire body pose determines positioning in exact time frame. Which means

that sequence of face vectors is a move, thus ability to detect human face vector allows us to determine the accuracy of the exercises, which in turn is useful for rehabilitation institutions, fitness centers, entertainment facilities.

In recent years, methods for recognizing the face, its individual parts and looks have evolved in several different directions. In order to recognize human movement in games, the Kinect optical system was developed, which has gained popularity due to its high recognition accuracy. But this system has significant restrictions on use, such as the maximum number of people in the image and the inability to use it in direct sunlight. In parallel, a different direction of recognition was developed from two-dimensional images without the use of depth sensors. This area is also divided into different methods, but in recent years, CNN-based methods have shown the best results in accuracy and speed of image analysis.

Different approaches to determining the vector of face rotation use different metrics. There are two main ways to express the position of the head: the position of the camera relative to the head, the deviation of the head from the position full face.

In a computer vision, position of the object means its relative orientation and camera position. You can change the pose by moving the subject relative to the camera or the camera relative to the subject. If we want to express the position of the head so that the goal is to find the pose of the object when we have a calibrated camera and we

know the location of n 3D points on the object and the corresponding 2D projections in the image.

A 3D hard object has only two types of camera movement. Translation – moving the camera from its current 3D location (X, Y, Z) to a new 3D location (X', Y', Z') is called a broadcast. As you can see, the translation has 3 degrees of freedom – you can move in the X, Y or Z direction. The translation is represented by a vector equal to $(X' - X, Y' - Y, Z' - Z)$. You can also rotate the camera around the X, Y , and Z axes. Therefore, rotation also has three degrees of freedom. You can represent it using the Euler angles (turn, step and slope), the 3×3 rotation matrix, or the direction of rotation (ie the axis) and angle.

To calculate the 3D position of an object in an image, you need the following information:

– 2D coordinates of multiple points. You need to position multiple points on the image (Fig. 1). In the case of the face, you can choose the corners of the eyes, the tip of the nose, the corners of the mouth, etc.

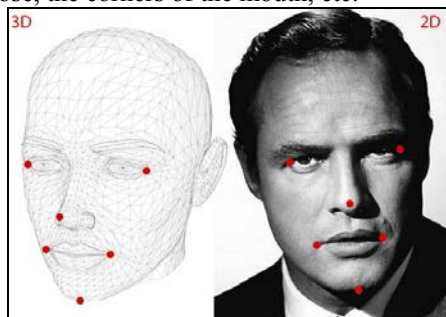


Figure 1 – Matching dots on 2D and 3D

– 3D locations of the same points. You need three-dimensional multi-point placements in an arbitrary frame. Since it is not possible to obtain an accurate three-dimensional model of any head from a single image, a generalized human head model is used.

Indicators that determine the position of the head when using the second way of expressing the position is the angle at which the head deviates from the frontal position in the following three directions: pitch, roll, yaw (Fig. 2). While the first method is suitable for determining the position of any object relative to the camera, the second method can only be used to determine the position of the head. Usually the second method is used in methods that are based on the creation of a CNN model.

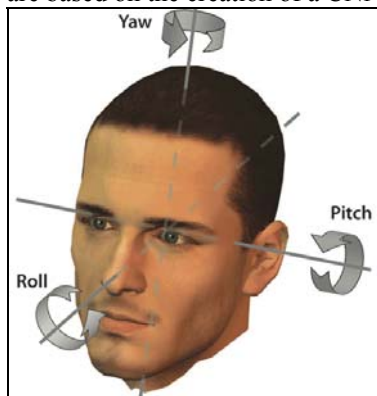


Figure 2 – Head rotation directions

One of the well-known technologies that play a crucial role in determining human postures and head postures is the Kinect camera developed by Microsoft. The Kinect camera has a clear advantage over other 3D cameras, because it gets more accurate depth information and works fast enough. Using Kinect, you can track up to six people at a time, as well as get motion analysis with an extraction function that allows you to determine the locations of the articulation points of the human skeleton. Extremely useful indoors, it cannot be used outdoors because the infrared depth sensor is extremely sensitive to sunlight. Methods based on the use of information obtained with the help of depth sensors allow to achieve accuracy in which the deviation from the real data is no more than two degrees [1].

The limitations on using Kinect, force us to look for ways to recognize a person's head based on a two-dimensional image of an RGB camera. Historically, there have been several basic approaches to face modeling: discriminatory, oriented approaches, and parameterized appearance models, or PAM. Regardless of the approach chosen, each method uses a face detector to obtain a region of the face image that will be further analyzed. There are several approaches to the problem of face recognition, some based on the implementation of the Viola-Jones algorithm [2], the wavelet transform [3], or the principal component method.

Considering the conducted researches and current restrictions **the purpose of this work** is creation of the method that will allow us to track human head moves and record only significant vectors of head direction to represent the entire move with minimal amount of data. **The object** of study is the vector of facial orientation, and **the subject** is the method for determining the reference points of head moves.

1 PROBLEM STATEMENT

Based on the study of existing systems and methods that perform the definition of the face vector on the images, we can describe a problem statement for the head tracking and comparison system.

As an input we have a sequence of frames with human head poses. In order to create a head tracking system, it is necessary to determine human head direction vector $\vec{a}_i, i = \overline{1, n}$ on each frame and distinguish those vectors $\vec{a}'_j, j = \overline{1, m}$ where head changes direction of the move. The output will be a sequence of reference points $\vec{a}'_1, \dots, \vec{a}'_m, m < n$ sufficient to describe the head moves.

The main criterion by which existing method should be considered is the accuracy of vector calculation and the stability of operation in different lighting conditions and video quality. The desired angle deviation shall be close to 2° for each rotation direction. The processing speed of one frame should be approximately 40 ms, in order to process high speed moves.

2 REVIEW OF THE LITERATURE

In recent years, methods that directly take 2D (face landmark) face points using deep learning tools appeared [4]. They have become dominant approaches to the analysis of face rotation due to their flexibility and resistance to extreme posture changes.

Recently, researches that use deep neural networks became leaders in the accuracy of the evaluation of head postures. J. Park and S. Kwon used deep neural networks such as Lenet to evaluate continuous head posture [5]. Massimiliano and Angelo explored the role of adaptive gradient methods for improving CNN performance in the evaluation of head postures [6]. The above works build their core unit using less than five collapsible layers and extract more expressive features from the training datasets. Kumar et al. modified the GoogleNet architecture to jointly predict facial landmarks and head postures. Xu et al. adapted the global and local CNN facial features for coarse-to-fine head posture evaluation. They used the global networks to predict the original head posture and the local networks to update the postures according to the current form. B. Huang et al. created a method of estimating head postures using two-stage groups with averaged top-k regression [7].

3 MATERIALS AND METHODS

Face landmark based methods that detect relative to camera head position are quite widespread and are commonly used with face-landmark detectors [8, 9]. The search of the pose is performed by determining the distortion applied to the 3D model. To achieve this, a landmark detector needs to find dozens of dots on the face, such as mouth corners, eye corners, jaw silhouettes, and more. Many algorithms have been developed and implemented in OpenCV. Identifying a person's landmarks begins with identifying the persons in the image and their extents (bounding box). The fastest and easiest way to detect faces in OpenCV is still to use the associated cascade classifiers using the `cv::CascadeClassifier` class provided in the core module. The tag detector will work around the detected individuals, starting from the bounding box.

Once we get the landmarks on the face, we can try to determine the direction of the face. 2D face orientations essentially correspond to the shape of the head. Therefore, in view of the three-dimensional model of the human head, we can find approximate corresponding three-dimensional points for a number of faces next way:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = s \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \end{pmatrix} \begin{pmatrix} U \\ V \\ W \\ 1 \end{pmatrix}. \quad (1)$$

In the above formula U, V, W are the position of the object in three-dimensional space, and x, y are the coordinates of the points in the image, f_x and f_y focal lengths in the x and y directions, c_x and c_y are focal center,

$r_i, i = \overline{1,9}$ are elements of the rotation vector, $t_i, i = \overline{1,3}$ are elements of camera translation vector. OpenCV provides implementation for finding rotation and moving with its `cv::solvePnP` functions `calib3d` module.

Since the general model is commonly used, this can lead to position recognition errors in case of significant differences between the test subjects' head and the model. Another problem in the application of this method depends on the accuracy of determining the anchor points of the individual. If a significant part of the points is absent, it is impossible to calculate the coordinates of the head.

The newest and more advanced methods of determining the posture of the head are methods based on the training of neural networks. Face detectors are trained on images of a person with different discrete poses and combine the outputs of a number of classifiers. These existing methods of estimating head postures are usually taught based on the classification of bin-poses or regression in one pose. Classification-based methods are performed by comparing the image with the labels of discrete poses, while regression-based methods directly output values of continuous posture. However, these methods are unlikely to use classification and regression loss at the same time. Classification-based methods are more robust to changes in non-ideal conditions. Their labels are true, accepting angular intervals (usually greater than 10°), so the corresponding labels are a bottleneck for further improving accuracy. Regression-based methods may more accurately predict head posture, but their effectiveness depends on the initial head position and variations in head posture, and perform poorly in non-ideal conditions.

Nonlinear regression methods use a training set to create nonlinear mapping from images to poses, and CNN is part of these methods. Because CNNs have the ability to reduce size and extract features automatically, they have achieved good results in various areas. The use of CNN has greatly improved the accuracy of head estimation, but excellent performance is only demonstrated in the same type of images and conditions that are found in the training set due to the intense overload of the training set.

FSA-Net [10] is representative of the regression method and uses a smooth stepped regression scheme. Existing methods of function aggregation process the input data as a set of features and thus ignore their spatial relation on the feature map. The method offers a fine-grained mapping of the structure for spatial grouping of features before aggregation.

FSA-Net performs spatial grouping of features before submitting them to the aggregation process. The developers of this method claim that due to this method they achieve accuracy that exceeds the capabilities of systems based on RGBD image processing, that is, images from cameras with depth sensors.

ResNet50 [11] combines a classification and regression approach to create a CNN that determines the

position of the head without first finding anchor points. The ResNet50 uses three separate losses, one for each corner. Each loss is a combination of two components: a classified posture and a regression component. Any network can be used and complemented by three fully connected layers that predict angles.

The idea behind this approach is that, using the classification, a softmax layer and cross-entropy are used, thus the network learns to clearly predict the adjacent posture. With three cross entropy losses, one for each Euler angle, we have three signals that are transmitted to the network to enhance learning. In order to get clearer forecasts, the expectation of each output angle for the resulting output class is calculated.

Then, regression losses are added to the network, namely the mean square error loss to improve the forecasts. There are three terminal losses, one at each angle, and each is a linear combination of both the corresponding classification and the regression losses.

In order to make progress in the prediction of image intensities, we need to find real datasets that contain accurate posture annotations, multiple images of different people, different lighting conditions, and a significant variety of poses. We identify two very different datasets that meet these requirements.

The BIWI dataset [12] is collected in the laboratory by recording RGB-D videos of different people with different head postures using a Kinect v2 device. It contains approximately 15,000 frames and rotation angles of $\pm 75^\circ$ for yaw, $\pm 60^\circ$ for pitch and $\pm 50^\circ$ for roll. This dataset is commonly used as a benchmark for estimating postures using depth methods, which confirms the accuracy of their labels. Examples of images from the BIWI dataset are shown in Fig. 3.

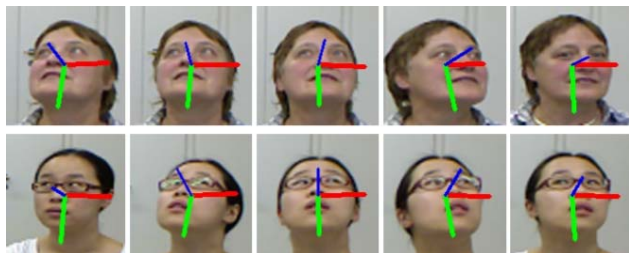


Figure 3 – An example from a BIWI dataset

Another 300W-LP data set [13] is a 300W extension that standardizes multiple alignment databases with 68 landmarks, including AFW, LFPW, HELEN, IBUG, and XM2VTS. This dataset offers synthetically advanced data to prepare landmark models. This is a collection of popular two-dimensional datasets that have been grouped and re-annotated. Synthetics of the set make it possible to create images with different values of deviation in all three directions and at the same time to receive the most accurate marks, which significantly affects the quality of the trained CNN models.

We propose to use head direction estimation methods to track head pose in real time. Considering the possible methods of comparison of head positions, two methods

were identified: frame-by-frame comparison, comparison of reference points. The first method is much simpler in terms of implementation, but it has significant drawbacks. First, the accuracy of this method depends on the speed of the face vector search method, if this speed is more than 40 ms (interval between frames when recording FPS 24 video), then some frames will be skipped, which will lead to a false negative scenario. Also, the person performing the exercises can change the pace of movement. Therefore, the fact of performing the exercise should not be associated with time.

We propose the reference point method that simplifies the process of recording the reference exercise. With this method we do not have to record every position of the head, but only those positions in which the movement changes direction. This method tracks the vector by which the face moves in time. Since the vector of the head can be represented as $\vec{a} = (r, p, y)$, where r is the angle of the head along the z axis, p is the angle of the head along the x axis, y is the angle of the head along the y axis, the reference point is a head direction vector in which the way it flows changes its direction. The reference point can be found by the following formula:

$$k < \left| \vec{a}_i - \vec{a}_{i+1} \right| \leq \left| \vec{a}_i - \vec{a}_{i-1} \right|, \quad (2)$$

where a_i is the value of the angle on the i frame and k is a flexibility degree. If the inequality holds for at least one of the metrics, the position is stored as a reference point. Flexibility degree is needed to avoid recording reference points caused by variations in vector estimation.

4 EXPERIMENTS

The methods discussed earlier were investigated by the following criteria:

- accuracy of the method;
- speed of work;
- maximal recognition angle.

The accuracy of the method of determining the position of the head is usually calculated in MAE (Mean absolute error). In statistics, the mean absolute error (MAE) is a measure of the difference between two continuous variables. Consider the scatter plot of n points where point i has coordinates (x_i, y_i) . Mean absolute error (MAE) is the average vertical distance between each point and an identical line. MAE is also the average horizontal distance between each point and an identical line. MAE is given by the following formula:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}. \quad (3)$$

Mean absolute error is the mean value of absolute errors $|e_i| = |y_i - x_i|$, where y_i is the predicted value and x_i is the real value.

The experiment was performed using two different datasets: BIWI, AFLW2000.

Speed was tested on frames obtained from a conventional camera with FPS 24. The calculations were performed using an Intel core i7 processor, RAM 16Gb.

The maximum recognition angle was investigated using a BIWI data set containing approximately 15,000 frames, and rotation angles $\pm 75^\circ$ for yaw, $\pm 60^\circ$ for pitch and $\pm 50^\circ$ for roll.

After the desired method was found we performed the experiment on flexibility degree to determine the most appropriate value that will help us to reduce amount of false reference points but not skip relevant points. Considered flexibility degrees were in interval between 1 and 2 degrees with step of 0.25.

5 RESULTS

Experiments mentioned above exposed next results. Comparisons of the main yaw, pitch, roll, and mean error values when using the BIWI data set are presented in figure 4.

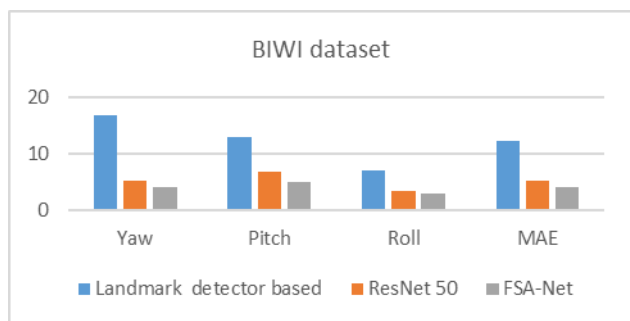


Figure 4 – The results experiment using BIWI dataset

The same experiment was performed with the AFLW2000 data set (Fig. 5). The main difference between these datasets is that BIWI is a laboratory-made data set, so the images have the same illumination and good quality. AFLW2000 [14] is a real-time data set where the brightness, blurring and image quality are very different.

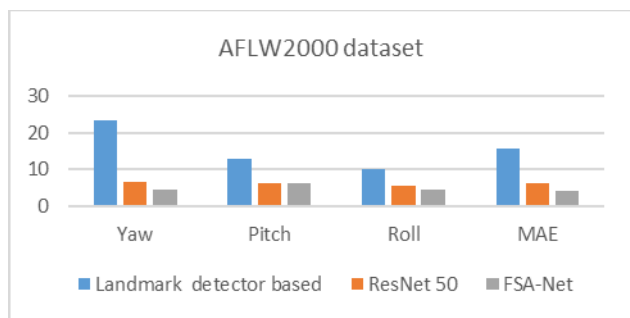


Figure 5 – The results experiment using AFLW2000 dataset

Speed test was performed under the same circumstances using same pictures for each method. The rate of calculation of postures for one frame and one face in the frame is shown in Table 1.

Table 1 – Speed test

№	Method	Average execution speed
1	Landmark-based detector	75 ms
2	ResNet 50	105 ms
3	FSA-Net	133 ms

The results of the accuracy of angle recognition depending on its value are shown in Fig. 6.

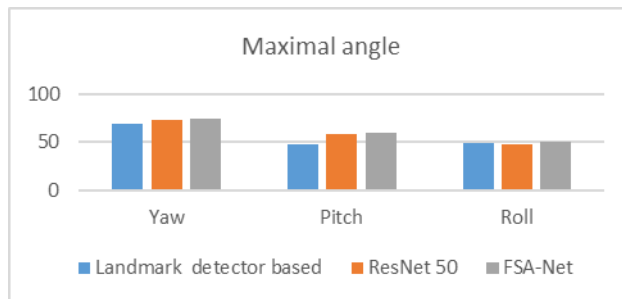


Figure 6 – Maximal angles of recognition

This coefficient was tested by next indicators: percent of correctly detected points, extra points.

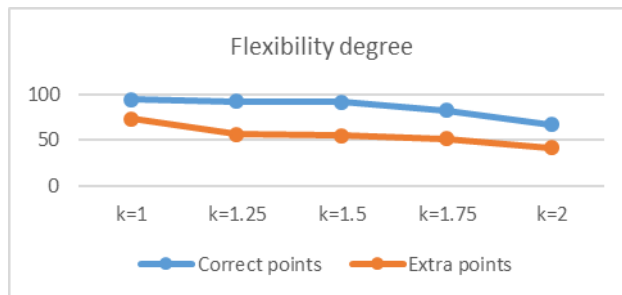


Figure 7 – Flexibility degree test results

6 DISCUSSIONS

The data obtained during the experiment shows that methods that are not based on the search for key points show quite good results, even on the AFLW2000 dataset. The average error for the landmark-based method is 15.8, which means an error of 15.8° . While ResNet 50 and FSA-Net allow for accuracy of 6.4° and 4.2° respectively. This leads us to believe that the potential of CNN-based methods is higher than key point methods.

Based on the results we have obtained during the speed test, the most accurate FSA-Net method shows the worst result of 133 ms, which is 3 times higher than the desired processing speed of 40 ms. If we use this method, we will be able to compare every fourth frame in the video. Such a low speed is due to the large number of image processing steps and is highly dependent on the speed of the face detector, which no method can do without.

All methods almost equally well recognize all available angles of the dataset. However, the landmark-based method showed the least accurate recognition of the difference between the boundary angles, which is due to the hidden part of the reference points when turning the head. For landmark-based methods, the lack of a large part of the points due to overlapping of the face with other

parts of the body or the headpiece makes it impossible to determine the angle.

According to the research, algorithms that are not based on facial landmark have shown much better accuracy results, especially on a real-time and poor lighting conditions. The 2° difference for the classification regression models is a very important step in improving the accuracy of face vector recognition. FSA-Net, which delivers 98.8% accuracy (based on an error of 4.2° from 360°), is the best candidate today for use in human head monitoring systems.

A method proposed to determine reference points shown that with flexibility degree close to 1°, 95% of all reference points are detected but too many extra points are recorded because of variations in head vector estimation. With flexibility degree equal to 1.25 or 1.5 accuracy stays the same while amount of extra points decreases. These values must be considered as most appropriate.

CONCLUSIONS

The practical value of the study is in creation of a reference points method. The proposed approach of search allows us to eliminate intermediate angles and store about 95% of all points where head move direction changes. This gives us ability to express head move with minimal amount of significant vectors and drastically reduces space needed to store the move representation. Further these reference points can be used to track human head moves by being compared to vectors determined in runtime. However further optimization is required by reference points search algorithm to decrease amount of extra points.

According to the results of the study, regression-based methods showed significantly better accuracy and independence of accuracy from light and partial face closure. All methods, and in particular the method based on facial landmark showed deterioration in the recognition of boundary angles due to the slight difference in the characteristics of the face. This is a window for the possible improvement of the method of operation in the conditions of tracking the person turned to the camera at an angle greater than 70°.

Significant optimization requires the speed of the FSA-Net method because at it works 2 times more slowly than landmark-based. Because the process of determining the posture of the head is multi-step, optimization can be performed both in the step of finding faces in the image, using faster detectors, and in the step of finding poses.

Research shows that methods of finding a person's head posture can compete with accuracy based on the use of depth sensors. However, methods of working with a simple RGB image have far fewer restrictions on the place of use.

ACKNOWLEDGEMENTS

The work is carried out in the framework of research directions of Software Engineering Department and with the support of researchers from the Scientific Research

Laboratory "Information Technologies in Learning and Computer Vision Systems" of Kharkiv National University of Radio Electronics.

REFERENCES

1. Borghi G., Fabbri M., Vezzani R. et al. ace-from-Depth for Head Pose Estimation on Depth Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, Vol. 42, pp. 596–609. DOI: 10.1109/TPAMI.2018.2885472
2. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features, *Computer Society Conference on Computer Vision and Pattern Recognition, 8–14 December 2001: proceedings*. Kauai, IEEE, 2001. DOI: 10.1109/CVPR.2001.990517
3. Shcherbakova G., Krylov V. N., Bilous N. V. Methods of automated classification based on wavelet-transform for automated medical diagnostics, *Information Technologies in Innovation Business Conference (ITIB), 7–9 October 2015: proceedings*. Kharkiv, IEEE, 2015, pp. 7–10. DOI: 10.1109/ITIB.2015.7355048
4. Wallhoff F., AblaBmeier M., Rigoll G. Multimodal Face Detection, Head Orientation and Eye Gaze Tracking, *International Conference on Multisensor Fusion and Integration for Intelligent Systems, 3–6 September 2006: proceedings*. Heidelberg, IEEE, 2006, pp. 13–18. DOI: 10.1109/MFI.2006.265612
5. Ahn B., Park J., Kweon I. S. Real-time head orientation from a monocular camera using deep neural network, *Asian Conference on Computer Vision: 12th Asian Conference on Computer Vision, 1–5 November 2014: proceedings*. Singapore, ACCV, 2014, pp. 82–96. DOI: 10.1007/978-3-319-16811-1_6
6. Patacchiola M., Cangelosi A. Head pose estimation in the wild using convolutional neural networks and adaptive gradient methods, *Pattern Recognition*, 2017, Vol. 71, pp. 132–143. DOI: 10.1016/j.patcog.2017.06.009
7. Huang B., Chen R., Xu W. et al Improving head pose estimation using two-stage ensembles with top-k regression, *Image and Vision Computing*, 2020, Vol. 93. DOI: 10.1016/j.imavis.2019.11.005
8. Kumar A., Alavi A., Chellappa R. KEPLER: Keypoint and pose estimation of unconstrained faces by learning efficient H-CNN regressors, *Image and Vision Computing*, 2018, pp. 258–265. DOI: 10.1016/j.imavis.2018.09.009
9. Hien L. T., Toan D. N., Lang T. V. Detection of Human Head Direction Based on Facial Normal Algorithm, *International Journal of Electronics Communication and Computer Engineering*, 2015, Vol. 6, pp. 110–114
10. Tsun-Yi Y., Yi-Ting C., Yen-Yu L. et al. FSA-Net: Learning Fine-Grained Structure Aggregation for Head Pose Estimation from a Single Image, *Conference on Computer Vision and Pattern Recognition (CVPR), 15–20 June 2019: proceedings*. Long Beach, IEEE, 2019, pp. 1087–1096. DOI: 10.1109/CVPR.2019.00118
11. Ruiz N., Chong E., Rehg J. M. Fine-grained head pose estimation without keypoints, *Conference on Computer Vision and Pattern Recognition Workshop, 18–22 June 2018: proceedings*. Salt Lake City, IEEE, 2018, pp. 1821–1829. DOI: 10.1109/CVPRW.2018.00281
12. Fanelli G., Dantone M., Gall J. et al. Random forests for real time 3D face analysis, *International Journal of Computer Vision*, 2013, Vol. 101, pp. 437–458. DOI: 10.1007/s11263-012-0549-0
13. Zhu X., Lei Z., Liu X. et al. Face alignment across large poses: A 3D solution, *Conference on Computer Vision and*

Pattern Recognition, 27–30 June 2016: proceedings. Las Vegas, 2016. – P. 146–155. DOI: 10.1109/CVPR.2016.23

14. Koestinger M., Wohlhart P., Roth P. M. et al. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization, *International Conference on Computer Vision Workshops*, 6–13

November 2011: proceedings. Barcelona, IEEE, 2011, pp. 2144–2151. DOI: 10.1109/ICCVW.2011.6130513

Received 15.05.2020.
Accepted 22.09.2020.

УДК 004.93

МЕТОД ОПОРНИХ ТОЧОК ДЛЯ ВІДСТЕЖЕННЯ РУХІВ ГОЛОВИ ЛЮДИНИ

Ракова А. О. – студент кафедри програмної інженерії, Харківський національний університет радіоелектроніки, Харків, Україна.

Білоус Н. В. – канд. техн. наук, доцент, професор кафедри програмної інженерії, Харківський національний університет радіоелектроніки, Харків, Україна.

АННОТАЦІЯ

Актуальність. Напрямок вектору обличчя людини є показником уваги людини. У нашому повсякденному житті він має багато застосувань, такі як взаємодія людина-комп'ютер, телеконференції, віртуальна реальність та 3D-передача звуку. Більше того, визначення положення голови можна використати для порівняння вправ, які виконує людина, з певним стандартом, що приводить нас до дослідження способів ефективного відстеження рухів. Системи на основі глибинних камер, які часто використовуються для цих цілей, мають суттєві недоліки, такі як зниження точності від прямого сонячного світла та необхідність додаткового обладнання. Розпізнавання від двовимірного зображення набуває все більшого поширення та усуває труднощі, пов'язані з глибинними камерами, що дозволяє використовувати їх у приміщенні та на відкритому повітрі.

Мета. Метою даної роботи є створення методу, який дозволить нам відстежувати рухи голови людини і записувати лише значні вектори напрямку голови.

Методи. У цій роботі пропонується метод опорних точок, який зменшує набір записаних векторів до мінімальної кількості, значущої для опису рухів голови. Він також досліджує та порівнює існуючі методи визначення вектору обличчя з точки зору використання у запропонованому підході.

Результати. Запропонований метод опорних точок показує здатність сильно зменшувати набір векторів напрямку голови, які описують рух. Відповідно до результатів дослідження, методи, засновані на регресії, показали значно кращу точність та незалежність від світла та часткового закриття обличчя, тому їх було обрано для використання в якості методів отримання вектору напрямку голови в підході опорних точок.

Висновки. Дослідження підтвердили застосовність методу опорних точок для відстеження рухів людини і показали, що методи визначення вектору голови людини за двовимірним зображенням можуть конкурувати в точності з методами на основі RGBD. Таким чином, у поєднанні із запропонованим підходом ці методи мають менше обмежень у використанні, ніж такі, що базуються на RGBD.

КЛЮЧОВІ СЛОВА: вектор направленості обличчя, рухи голови, розпізнавання, глибоке навчання.

УДК 004.93

МЕТОД ОПОРНЫХ ТОЧЕК ДЛЯ ОТСЛЕЖИВАНИЯ ДВИЖЕНИЙ ГОЛОВЫ ЧЕЛОВЕКА

Ракова А. О. – студент кафедры программной инженерии, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

Белоус Н. В. – канд. техн. наук, доцент, профессор кафедры программной инженерии, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

АННОТАЦИЯ

Актуальность. Направление вектора человеческого лица является индикатором человеческого внимания. Оно имеет множество вариантов применения в нашей повседневной жизни, таких как взаимодействие человека с компьютером, телеконференций, виртуальной реальности и 3D-рендеринга звука. Более того, определение положения головы можно использовать для сравнения упражнений, выполняемых человеком с определенным стандартом, что приводит нас к исследованию способов эффективного отслеживания движений. Системы на основе глубинных камер, часто используемые для этих целей, имеют существенные недостатки, такие как снижение точности на прямом солнечном свете и необходимость дополнительного оборудования. Распознавание по двумерному изображению становится все более распространенным и устраняет трудности, связанные с глубинными камерами, что позволяет использовать их в помещении и на улице.

Цель. Целью данной работы является создание метода, который позволит нам отслеживать движения головы человека и фиксировать только значимые векторы направления головы.

Методы. Эта статья предлагает метод опорных точек, который уменьшает набор записанных векторов до минимального значения, значимого для описания движений головы. Он также исследует и сравнивает существующие методы определения вектора лица с точки зрения использования в предлагаемом подходе.

Результаты. Предложенный метод опорных точек показывает способность значительно уменьшить набор векторов направления головы, которые описывают движение. Согласно результатам исследования, методы на основе регрессии

показали значительно лучшую точность и независимость от легкого и частичного закрытия лица, поэтому они были выбраны для использования в качестве методов для получения вектора направления головы в методе опорных точек.

Выводы. Исследования подтвердили применимость метода опорных точек для отслеживания движений человека и показали, что методы определения вектора головы человека по двумерному изображению могут конкурировать в точности с методами на основе RGBD. Таким образом, в сочетании с предлагаемым подходом, эти методы создают меньше ограничений в использовании, чем основанные на RGBD.

КЛЮЧЕВЫЕ СЛОВА: вектор направленности лица, движение головы, распознавание, глубокое обучение.

ЛІТЕРАТУРА / LITERATURA

1. Face-from-Depth for Head Pose Estimation on Depth Images / [G. Borghi, M. Fabbri, R. Vezzani et al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2018. – Vol. 42. – P. 596–609. DOI: 10.1109/TPAMI.2018.2885472
2. Viola P. Rapid object detection using a boosted cascade of simple features / P. Viola, M. Jones // Computer Society Conference on Computer Vision and Pattern Recognition, 8–14 December 2001: proceedings. – Kauai : IEEE, 2001. DOI: 10.1109/CVPR.2001.990517
3. Shcherbakova G. Methods of automated classification based on wavelet-transform for automated medical diagnostics / G. Y. Shcherbakova, V. N. Krylov, N. V. Bilous // Information Technologies in Innovation Business Conference (ITIB), 7–9 October 2015: proceedings. – Kharkiv: IEEE, 2015. – P. 7–10. DOI: 10.1109/ITIB.2015.7355048
4. Wallhoff F. Multimodal Face Detection, Head Orientation and Eye Gaze Tracking / F. Wallhoff, M. AblaBmeier, G. Rigoll // International Conference on Multisensor Fusion and Integration for Intelligent Systems, 3–6 September 2006: proceedings. – Heidelberg: IEEE, 2006. – P. 13–18. DOI: 10.1109/MFI.2006.265612
5. Ahn B. Real-time head orientation from a monocular camera using deep neural network / B. Ahn, J. Park, I. S. Kweon // Asian Conference on Computer Vision: 12th Asian Conference on Computer Vision, 1–5 November 2014: proceedings. – Singapore: ACCV, 2014. – P. 82–96. DOI: 10.1007/978-3-319-16811-1_6
6. Patacchiola M. Head pose estimation in the wild using convolutional neural networks and adaptive gradient methods / M. Patacchiola, A. Cangelosi // Pattern Recognition. – 2017. – Vol. 71. – P. 132–143. DOI: 10.1016/j.patcog.2017.06.009
7. Improving head pose estimation using two-stage ensembles with top-k regression / [B. Huang, R. Chen, W. Xu et al.] // Image and Vision Computing. – 2020. – Vol. 93. DOI: 10.1016/j.imavis.2019.11.005
8. Kumar A. KEPLER: Keypoint and pose estimation of unconstrained faces by learning efficient H-CNN regressors / A. Kumar, A. Alavi, R. Chellappa // Image and Vision Computing. – 2018. – P. 258–265. DOI: 10.1016/j.imavis.2018.09.009
9. Hien L.T. Detection of Human Head Direction Based on Facial Normal Algorithm / L. T. Hien, D. N. Toan, T. V. Lang // International Journal of Electronics Communication and Computer Engineering. – 2015. – Vol. 6. – P. 110–114
10. FSA-Net: Learning Fine-Grained Structure Aggregation for Head Pose Estimation from a Single Image / [Y. Tsun-Yi, C. Yi-Ting, L. Yen-Yu et al.] // Conference on Computer Vision and Pattern Recognition (CVPR), 15–20 June 2019: proceedings. – Long Beach: IEEE, 2019. – P. 1087–1096. DOI: 10.1109/CVPR.2019.00118
11. Ruiz N. Fine-grained head pose estimation without keypoints / N. Ruiz, E. Chong, J. M. Rehg // Conference on Computer Vision and Pattern Recognition Workshop, 18–22 June 2018: proceedings. – Salt Lake City : IEEE, 2018. – P. 1821–1829. DOI: 10.1109/CVPRW.2018.00281
12. Random forests for real time 3D face analysis / [G. Fanelli, M. Dantone, J. Gall et al.] // International Journal of Computer Vision. – 2013. – Vol. 101, P. 437–458. DOI: 10.1007/s11263-012-0549-0
13. Face alignment across large poses: A 3D solution / [X. Zhu, Z. Lei, X. Liu et al.] // Conference on Computer Vision and Pattern Recognition, 27–30 June 2016: proceedings. – Las Vegas, 2016. – P. 146–155. DOI: 10.1109/CVPR.2016.23
14. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization / [M. Koestinger, P. Wohlhart, P. M. Roth et al.] // International Conference on Computer Vision Workshops, 6–13 November 2011: proceedings. – Barcelona : IEEE, 2011. – P. 2144–2151. DOI: 10.1109/ICCVW.2011.6130513

ОЦЕНКА ИНФОРМАТИВНОСТИ И ОТБОР ЭКЗЕМПЛЯРОВ НА ОСНОВЕ ХЭШИРОВАНИЯ

Субботин С. А. – д-р техн. наук, профессор, заведующий кафедрой программных средств Национального университета «Запорожская политехника», Запорожье, Украина.

АННОТАЦИЯ

Актуальность. Для сокращения размерности данных при построении диагностических и распознающих моделей возникает необходимость отбора наиболее информативных экземпляров, а также отбора наиболее информативных признаков. Затраты времени на отдельную реализацию данных процедур являются высокими вследствие итеративности и взаимосвязанности данных процедур.

Цель. Цель работы – сокращение временных затрат на сокращение размерности данных путем создания метода отбора наиболее информативных экземпляров на основе хэширования.

Метод. Предложен метод расчета весов для определения хэшей экземпляров, который детерминированным способом определяет веса признаков на основе их рангов, которые, в свою очередь, определяет с учетом числа равных разбиений диапазонов признаков, минимально достаточного для выделения кластеров на оси признака с приемлемой точностью. Это позволяет исключить необходимость итеративного перебора различных комбинаций признаков, определения случайных проекций признаков, а также решение итеративных оптимизационных задач поиска наилучшей проекции признаков, что существенно сокращает затраты времени на расчет весов, при этом обеспечивая локальную чувствительность хэша. Полученные хэши возможно использовать как для отбора экземпляров, так и для отбора признаков.

Предложен метод определения индивидуальной и групповой значимости экземпляров выборки, в котором использует как меру сходства расстояния между хэшами экземпляров и по аналогии с методом потенциалов находит потенциалы, наводимые классами на каждый экземпляр, а на их основе определяет показатели значимости экземпляров, исходя из того, что экземпляр в пространстве признаков тем информативнее, чем меньше минимальная разность потенциалов классов, наводимых на экземпляр.

Предложен метод определения оценок информативности признаков, который на основе нормирования весов, полученных при формировании хэшей, определяет показатели информативности признаков, отдавая предпочтение признакам с меньшим числом разбиений.

Результаты. Проведено экспериментальное исследование, подтвердившее работоспособность предложенных методов при решении практических задач.

Выводы. Разработанное математическое обеспечение может быть рекомендовано для решения задач сокращения размерности данных.

КЛЮЧЕВЫЕ СЛОВА: экземпляр, признак, информативность, хэширование, хэш, сокращение размерности выборки.

НОМЕНКЛАТУРА

δ – заданная пользователем константа, регулирующая допустимое расхождение значений критериев качества редуцированной и исходной выборок;

ε – максимально допустимое значение ошибки;

$d_{*}^{s,p}$ – расстояние хэшей s -го и p -го экземпляров;

$E_{j,q}$ – ошибка для каждого q -го интервала значений j -го признака;

E_j – суммарная ошибка для всех интервалов j -го признака;

F' – критерий качества полученной редуцированной выборки;

F – критерий качества исходной выборки;

G – группа экземпляров;

I_{*}^s – показатель индивидуальной значимости s -го экземпляра;

$I_{*}(G)$ – оценки групповой информативности экземпляров в группе G ;

I_j – показатели индивидуальной информативности признаков;

j – номер экземпляра выборки;

k – номер класса;

K – число классов;

M – объем памяти ЭВМ, использованной при построении модели на основе исходной выборки;

M' – объем памяти ЭВМ, использованной при построении модели на основе редуцированной выборки;

N – число признаков, характеризующих экземпляры редуцированной выборки;

N' – число признаков, характеризующих экземпляры выборки;

n – размерность выборки;

n' – размерность редуцированной выборки;

$P^k(x_{*}^s)$ – потенциал, наводимый экземплярами разных классов на хэш s -го экземпляра;

Q – число равных по длине интервалов, на которые разбиваются диапазоны значений признаков;

Q_j – число интервалов, на которые разбивается диапазон значений j -го признака;

q – номер интервала значений признака;

r_j – ранг j -го признака;

$S_k^{j,q}$ – число экземпляров k -го класса, попавших в него;

S' – число экземпляров в редуцированной выборке;
 s – номер экземпляра;
 S – число экземпляров в выборке;
 t – время построения модели на основе исходной выборки;
 t' – время построения модели на основе редуцированной выборки;
 w_j – вес j -го признака;
 x – набор экземпляров исходной выборки;
 x' – набор экземпляров редуцированной выборки;
 x^s – s -й экземпляр выборки;
 x_j^s – значение j -го входного признака, сопоставленное s -му экземпляру выборки;
 x_j^* – хэш s -го экземпляра выборки;
 y – набор значений исходного признака;
 y' – набор значений выходного признака, сопоставленных экземплярам редуцированной выборки;
 y^s – значение выходного признака, сопоставленное s -му экземпляру выборки.

ВВЕДЕНИЕ

В задачах построения диагностических и распознающих моделей по прецедентам зачастую приходится сталкиваться с проблемой большой размерности данных [1], когда число экземпляров является довольно большим, а число характеризующих их признаков также велико. Построение моделей в таких задачах, как правило, сопряжено с большими затратами машинного времени и памяти ЭВМ. Поэтому возникает необходимость сокращения размерности данных [2].

Объектом исследования являлся процесс сокращения размерности данных.

Сокращение размерности данных возможно осуществить путем отбора наиболее информативных признаков [1, 3], путем отбора наиболее значимых экземпляров [3–9], а также путем формирования искусственных признаков [10, 11].

Отбор информативных признаков (feature selection) [1, 3] является наиболее хорошо разработанным и широко используемым инструментом для сокращения размерности данных. Однако он эффективен только для ситуаций, когда среди исходных признаков имеется достаточно информативных признаков. Часто же на практике признаки могут быть индивидуально мало информативными. Кроме того, отбор признаков требует, как правило, значительных затрат времени на перебор комбинаций признаков. Временные затраты на отбор признаков существенно зависят от числа используемых экземпляров.

Выделение искусственных признаков (feature extraction) [10, 11] является узко применимым инструментом, пригодным в основном для распознавания изображений. Создание универсального для всех областей и вместе с тем эффективного метода формирования искусственных признаков представляется не-

решенной и практически нереализуемой в настоящее время задачей.

Отбор наиболее значимых экземпляров (instance selection) [3–10] является менее широко используемым инструментом на практике по сравнению с отбором информативных признаков, однако он позволяет существенно сократить затраты времени на построение распознающей или диагностической модели, поскольку позволяет исключить из обучающей выборки незначимые для построения модели наблюдения, что еще до начала обучения модели обеспечивает существенное повышение ее обобщающих свойств по сравнению с использованием всей исходной выборки наблюдений.

Предметом исследования являлись методы отбора информативных экземпляров для построения диагностических и распознающих моделей.

При отборе экземпляров для формирования обучающих выборок приходится оперировать исходным набором признаков, что существенно влияет на затраты времени на отбор экземпляров. Поэтому возникает необходимость сжатия описания экземпляров, но без решения задачи отбора информативных признаков. Одним из возможных путей решения этой задачи является использование хэширования [12–14].

Целью работы являлось сокращение временных затрат на сокращение размерности данных путем создания метода отбора наиболее информативных экземпляров на основе хэширования.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть задана исходная выборка наблюдений $\langle x, y \rangle$, $x = \{x^s\}$, $x^s = \{x_j^s\}$, $y = \{y^s\}$, $s = 1, 2, \dots, S$, $j = 1, 2, \dots, N$.

Тогда задача сокращения размерности выборки $\langle x, y \rangle$ состоит в том, чтобы получить $\langle x', y' \rangle$: $x' \subseteq x$, $y' \subseteq y$, $S' \leq S$, $N' \leq N$. При этом критерий качества полученной редуцированной выборки F' должен принимать приемлемое значение относительно значения критерия качества для исходной выборки F : $|F - F'| \leq \delta$.

Данная задача может быть разбита на задачу отбора наиболее информативных признаков, задачу отбора наиболее информативных экземпляров и задачу формирования искусственных признаков.

Задача отбора наиболее значимых экземпляров состоит в том, чтобы для выборки $\langle x, y \rangle$ получить $\langle x', y' \rangle$: $x' \subseteq x$, $y' \subseteq y$, $S' \leq S$, $N' = N$.

Задача отбора наиболее значимых признаков состоит в том, чтобы для выборки $\langle x, y \rangle$ получить $\langle x', y' \rangle$: $x' \subseteq x$, $y' = y$, $S' = S$, $N' \leq N$.

Задача формирования искусственных признаков для выборки $\langle x, y \rangle$ состоит в том, чтобы получить $\langle x', y' \rangle$: $x' = f(x)$, $y' = y$, $S' = S$, $N' \leq N$.

Для заданной выборки наблюдений $\langle x, y \rangle$ задача формирования хэширующего преобразования состоит в том, чтобы получить $\langle x', y' \rangle$: $x' = f(x)$, $y' = y$, $S' = S$, $N' = 1$.

2 ОБЗОР ЛИТЕРАТУРЫ

Методы отбора экземпляров можно разделить на детерминированные [3], которые основаны на жестко заданных стратегиях поиска решений и стохастические [5], которые представляют собой стратегии перебора случайно формируемых наборов решений. Обе группы методов отбора экземпляров требуют задания критерия отбора и оценки решений, критерия останова, а также способов формирования новых решений на основе ранее рассмотренных.

Хэширующие преобразования [12–14] отображают экземпляры из N -мерного исходного пространства признаков на одномерную ось хэша. По сути хэширование можно рассматривать как разновидность методов формирования искусственных признаков.

Ключевым свойством хэшей для задач распознавания образов является сохранение пространственной топологии исходного пространства признаков на оси хэша – признака. Поэтому среди всех известных методов расчета хэшей целесообразно ограничиться локально чувствительным хэшированием [15–17], которое стремится построить хэширующие преобразования, которые позволят отобразить расстояния между экземплярами в исходном пространстве признаков в расстояния в пространстве.

Многие методы локально чувствительного хэширования [15–17] определяют хэш для s -го экземпляра выборки по формуле:

$$x_*^s = \sum_{j=1}^N w_j x_j^s.$$

Именно способ расчета весов признаков определяет отличие методов данной группы. В большинстве известных методов веса определяются в результате итеративного перебора случайных отображений из исходного набора признаков на ось хэша, что является весьма затратным по времени. Поэтому для ускорения процесса определения весов необходимо разработать детерминированный метод.

Определив хэши экземпляров, необходимо на их основе создать показатели, позволяющие оценивать значимость экземпляров в пространстве хэша, что позволит исключить необходимость загрузки в память ЭВМ многомерной выборки, а также оперирования многомерными описаниями экземпляров. Это позволит также значительно сократить объем вычислений по сравнению с обработкой выборки во всем исходном пространстве признаков.

3 МАТЕРИАЛЫ И МЕТОДЫ

Для исключения перебора случайных проекций выборки из исходного пространства будем рассматривать иерархию разбиений пространства признаков на области, заменяя в общем случае вещественные значения признаков на дискретные номера интервалов по оси признака, стремясь для каждого признака найти такое разбиение на интервалы, при котором число

© Субботин С. А., 2020
 DOI 10.15588/1607-3274-2020-3-12

интервалов будет наименьшим, но обеспечивающим требуемую точность. Тогда веса признаков определим с учетом числа интервалов, сформированных для каждого признака. Чем меньше нужно интервалов для обеспечения приемлемой точности, тем более ценным является соответствующий признак.

Формально предложенный метод хэширования, реализующий описанные выше идеи, можно представить следующим образом.

Этап Инициализации. Задать исходную выборку $\langle x, y \rangle$, а также максимально допустимое значение ошибки $0 \leq \varepsilon \ll S$. Нормировать значения признаков, отобразив их на интервал $[0, 1]$:

$$x_j^s = \frac{x_j^s - \min_{i=1,2,\dots,N} \{x_i^s\}}{\max_{i=1,2,\dots,N} \{x_i^s\} - \min_{i=1,2,\dots,N} \{x_i^s\}}.$$

Этап задания предела разбиения признаков. Определить предельное число равных по длине интервалов Q , на которые разбиваются диапазоны значений признаков: не более чем S , но не менее чем K . Эвристически можно рекомендовать принять:

$$Q = S \\ \text{или} \\ Q = \max\{\lceil \log_2 S \rceil, \min\{S, 2K\}\}.$$

Для всех $j = 1, 2, \dots, N$ принять число интервалов, на которые разбивается диапазон значений j -го признака, $Q_j = Q$.

Этап разбиения признаков. Для каждого j -го признака, $j = 1, 2, \dots, N$ выполнить последовательно пункты 1–4:

1. Разбить диапазон значений j -го признака на Q_j равных по длине интервалов.

2. Для каждого q -го интервала значений j -го признака, $q = 1, 2, \dots, Q_j$, определить:

– число экземпляров k -го класса, попавших в него $S_k^{j,q}$, $k = 1, 2, \dots, K$;

– ошибку для каждого q -го интервала значений j -го признака:

$$E_{j,q} = \sum_{s=1}^S \sum_{p=s+1}^S \{ | -1 \leq Qx_j^s - q \leq 0, -1 \leq Qx_j^p - q \leq 0, y^s \neq y^p \}.$$

3. Определить суммарную ошибку для всех интервалов j -го признака:

$$E_j = \sum_{q=1}^{Q_j} E_{j,q}.$$

4. Если ошибка для всех интервалов j -го признака E_j является приемлемой ($E_j \leq \varepsilon$), то сократить число интервалов j -го признака в два раза, установив $Q_j = \lceil Q_j / 2 \rceil$, и перейти к п. 1, в противном случае – вернуть предыдущее разбиение диапазона j -го признака.

Этап ранжирования признаков. Для $j = 1, 2, \dots, N$ определить ранги признаков r_j в порядке увеличения Q_j : чем больше Q_j , тем меньше ранг j -го признака. Для признаков с одинаковыми значениями Q_j эвристически считать более важным (с большим рангом) тот, который индивидуально более значим для выходной переменной, или признак с меньшим номером.

Этап расчета весов. Установить веса признаков:

$$w_j = \left(\max_{j=1,2,\dots,N} \{Q_j\} \right)^{N-r_j},$$

либо

$$w_j = \left(2^{\left\lceil \log_2 \max_{j=1,2,\dots,N} \{Q_j\} \right\rceil} \right)^{N-r_j}.$$

В результате выполнения предложенного метода будет получен набор весов, позволяющий определять локально чувствительные хэши экземпляров.

Хэши экземпляров, полученных на основе весов признаков, определенных предложенным выше методом, можно использовать для оценки информативности и отбора экземпляров. Побочным результатом предложенного метода являются веса признаков, которые возможно использовать не только для определения хэшей, но также и для оценивания информативности признаков.

Поскольку хэши, рассчитанные на основе предложенного метода, являются аналогом расстояния между экземплярами, то для них подобно методу потенциалов [18] возможно определить потенциалы, наводимые классами. В свою очередь, показатель значимости экземпляра возможно определить на основе сопоставления потенциалов классов, наводимых на экземпляр. Комбинируя определенным образом показатели индивидуальной значимости экземпляров целесообразно определить показатели групповой значимости экземпляров.

Метод оценивания значимости экземпляров, реализующий описанные выше идеи, возможно представить следующим образом.

Этап инициализации. Задать исходную выборку $\langle x, y \rangle$ и нормировать её.

Этап определения весов признаков. Используя предложенный выше метод, найти веса признаков для расчета хэшей.

Этап определения хэшей. Для экземпляров выборки найти хэши $\{x_*^s\}$, $s = 1, 2, \dots, S$, используя полученные веса.

Этап расчета потенциалов. Для каждого s -го хэша экземпляра, $s = 1, 2, \dots, S$, определить потенциал, наводимый экземплярами разных классов на данный экземпляр:

$$P^k(x_*^s) = \sum_{p=1}^S \left\{ \frac{1}{1 + d_*^{s,p}} \mid p \neq s, y^p = k \right\},$$

$$k = 1, 2, \dots, K, s = 1, 2, \dots, S.$$

С неявно заданным учетом весов признаков определим:

$$d_*^{s,p} = \delta_*^{s,p},$$

$$\delta_*^{s,p} = \left| x_*^s - x_*^p \right|.$$

Без учета весов признаков определим расстояние хэшей следующим образом:

$$d_*^{s,p} = \sum_{j=1}^N \left((\delta_*^{s,p} \bmod w_{\arg \min_{m=1,2,\dots,N} \{r_m | r_m - r_j = 1\}}) - (\delta_*^{s,p} \bmod w_j) \right).$$

Этап определения оценок индивидуальной значимости экземпляров.

Экземпляр в пространстве признаков тем легче отделить от других экземпляров, чем больше минимальная разность потенциалов классов, наводимых на него. То есть, чем меньше минимальная разность потенциалов классов, наводимых на экземпляр, тем сложнее его отделить от экземпляров в пространстве признаков – такой экземпляр будет ценнее по сравнению с другими экземплярами для построения модели, т.к. он вероятно ближе к межклассовой границе.

Соответственно, определим показатель индивидуальной значимости s -го экземпляра (стратегия минимума разности потенциалов):

$$I_*^s = \frac{1}{1 + \min_{k=1,\dots,K} \{ \min_{q=k+1,\dots,K} \{ |P^k(x_*^s) - P^q(x_*^s)| \} \}}.$$

Данный показатель будет принимать значения от нуля до единицы. Он будет равен единице в случае, когда потенциал одного из классов, наводимый на экземпляр, не будет отличаться от потенциала другого класса, наводимого на этот же экземпляр. Чем сильнее будет удаленность экземпляра от межклассовой границы, тем меньше будет значение данного показателя.

Альтернативно показатель индивидуальной значимости s -го экземпляра возможно определить на основе среднего значения разностей потенциалов:

$$I_*^s = \frac{1}{1 + \frac{1}{0,5K(K-1)} \sum_{k=1}^K \sum_{q=k+1}^K |P^k(x_*^s) - P^q(x_*^s)|}.$$

Данный показатель будет принимать значения от нуля до единицы. Он будет равен единице в случае, когда в среднем потенциалы классов, наводимые на экземпляр, не будут отличаться от потенциалов других классов, наводимых на этот же экземпляр. Чем сильнее будет удаленность экземпляра от межклассовой границы, тем меньше будет значение данного показателя.

Аналогичным образом, комбинируя минимальную и среднюю стратегии объединения разностей потенциалов, наводимых разными классами на данный экземпляр, определим показатели:

– на основе стратегии минимума средней разности потенциалов:

$$I_*^s = \frac{1}{1 + \min_{k=1, \dots, K} \left\{ \frac{1}{(K-k)} \sum_{q=k+1}^K |P^k(x_*^s) - P^q(x_*^s)| \right\}}$$

– на основе стратегии среднего минимума разности потенциалов:

$$I_*^s = \frac{1}{1 + \frac{1}{K} \sum_{k=1}^K \min_{q=k+1, \dots, K} \{|P^k(x_*^s) - P^q(x_*^s)|\}}$$

Этап определения групповых оценок информативности экземпляров. На основе рассчитанных индивидуальных оценок информативности экземпляров определить оценки групповой информативности экземпляров на основе одной из стратегий:

– на основе стратегии минимума индивидуальных оценок информативности экземпляров в группе G :

$$I_*(G) = \min_{s=1, 2, \dots, S} \{I_*^s | x^s \in G\};$$

– на основе стратегии средней индивидуальной информативности экземпляров в группе G :

$$I_*(G) = \frac{1}{S} \sum_{s=1}^S \{I_*^s | x^s \in G\}.$$

Индивидуальные и групповые оценки значимости экземпляров выборки, полученные на основе предложенного метода, могут бы использованы как в детерминированных та и в стохастических методах формирования выборок. Важной особенностью полученных показателей значимости экземпляров является то, что они не требуют построения моделей для определения значимости экземпляров.

Для оценивания информативности признаков на основе весов, полученных при формировании хэшей, пронормируем веса, получив таким образом показатель индивидуальной информативности признаков:

$$I_j = \frac{w_j}{\sum_{i=1}^N w_i}$$

либо

$$I_j = \frac{w_j}{\max_{i=1, 2, \dots, N} \{w_i\}},$$

либо

$$I_j = \frac{w_j - \min_{i=1, 2, \dots, N} \{w_i\}}{\max_{i=1, 2, \dots, N} \{w_i\} - \min_{i=1, 2, \dots, N} \{w_i\}}.$$

Предложенный показатель индивидуальной информативности признаков будет принимать значения от нуля до единицы: чем меньше будет его значение, тем менее информативным является признак, чем больше будет его значение, тем более индивидуально информативным является признак.

4 ЭКСПЕРИМЕНТЫ

Для изучения свойств предложенных методов они были программно реализованы и исследованы путем решения ряда практических задач распознавания и диагностирования [19–21], характеристики которых приведены в табл. 1. Здесь размерность выборки $n = NS$.

Для каждой практической задачи проводились эксперименты по сокращению размерности выборки до достижения заданной приемлемой точности – отдельно на основе отбора экземпляров, отдельно на основе отбора признаков, а также совместно на основе отбора экземпляров и признаков. Модели строились на основе многослойной нейронной сети прямого распространения сигнала. Число слоев последовательно менялось от одного до трех. Число узлов входного слоя задавалось равным числу используемых признаков, на последнем слое располагался один выходной нейрон (задачи рассматривались как бинарные). Число нейронов скрытого слоя подбиралось в автоматическом режиме там, чтобы при минимальном числе нейронов обеспечить приемлемую точность.

5 РЕЗУЛЬТАТЫ

Результаты проведенных экспериментов приведены в табл. 2.

Таблица 1 – Характеристики практических задач

Задача	Описание	N	S	n	K
Iris	Классификация ирисов Фишера [19]	4	150	600	2
Plant	Классификация сельскохозяйственных растений по данным дистанционного зондирования [20]	55	248	13640	2
Bronch	Дифференциальная диагностика хронического обструктивного бронхита [21]	28	205	5740	2
Vehicle	Распознавание типа автотранспортного средства по изображению [22]	26	3992	103792	2

Таблица 2 – Результаты экспериментов

Задача	Тип эксперимента	S/S'	N/N'	n/n'	t/t'	M/M'
Iris	Отбор экземпляров	1,88	1,00	1,88	1,76	1,44
Plant	Отбор экземпляров	3,40	1,00	3,44	2,97	2,36
Plant	Отбор признаков	1,00	6,11	6,11	4,58	5,33
Plant	Отбор экземпляров и признаков	3,31	6,11	18,19	7,82	11,71
Bronch	Отбор экземпляров	2,01	1,00	2,03	1,57	1,51
Bronch	Отбор признаков	1,00	1,56	1,65	1,15	1,56
Bronch	Отбор экземпляров и признаков	1,92	1,47	2,82	2,78	2,04
Vehicle	Отбор экземпляров	2,18	1,00	2,18	1,66	1,59
Vehicle	Отбор признаков	1,00	1,73	1,73	1,35	1,40
Vehicle	Отбор экземпляров и признаков	2,16	1,53	3,12	2,90	2,38

6 ОБСУЖДЕНИЕ

Как видно из табл. 2, выборки, редуцированные с помощью предложенных методов, позволяют достигать приемлемой точности при существенном сокращении числа экземпляров и числа признаков, а сами хэши могут использоваться как замена или дополнение исходного или сокращенного набора признаков при построении диагностических и распознающих моделей. Использование редуцированных выборок позволяет сократить затраты времени и памяти на построение моделей, а также повышает их обобщающие свойства по сравнению с моделями, обучаемыми на основе исходных выборок.

По сравнению с методами локально чувствительного хэширования [15–17] предложенный метод определения весов не требует итеративного перебора преобразований и является детерминированным.

По сравнению со стохастическими методами формирования выборок [5, 8] предложенный метод расчета показателей информативности экземпляров не требует построения моделей, а также является детерминированным. По сравнению с детерминированными методами [4] предложенный метод позволяет в автоматическом режиме оценить значимость экземпляров без участия человека, а также с учетом топологии классов.

ВЫВОДЫ

Решена актуальная задача создания метода отбора наиболее информативных экземпляров на основе хэширования для уменьшения временных затрат на сокращение размерности данных.

Научная новизна полученных результатов состоит в том, что:

– предложен метод расчета весов для определения хэшей экземпляров, который детерминированным способом определяет веса признаков на основе их рангов, которые, в свою очередь, определяет с учетом числа равных разбиений диапазонов признаков, минимально достаточного для выделения кластеров на оси признака с приемлемой точностью. Это позволяет исключить необходимость итеративного перебора различных комбинаций признаков, определения случайных проекций признаков, а также решение итеративных оптимизационных задач поиска наилучшей проекции признаков, что существенно сокращает за-

траты времени на расчет весов, при этом обеспечивая локальную чувствительность хэша. Полученные хэши возможно использовать как для отбора экземпляров, так и для отбора признаков;

– предложен метод определения индивидуальной и групповой значимости экземпляров выборки, в котором использует как меру сходства расстояния между хэшами экземпляров и по аналогии с методом потенциалов находит потенциалы, наводимые классами на каждый экземпляр, а на их основе определяет показатели значимости экземпляров, исходя из того, что экземпляр в пространстве признаков тем информативнее, чем меньше минимальная разность потенциалов классов, наводимых на экземпляр;

– предложен метод определения оценок информативности признаков, который на основе нормирования весов, полученных при формировании хэшей, определяет показатели информативности признаков, отдавая предпочтение признакам с меньшим числом разбиений.

Практическая ценность полученных результатов состоит в том, что проведено экспериментальное исследование, подтвердившее работоспособность предложенных методов при решении практических задач распознавания и диагностирования. Разработанное математическое обеспечение может быть рекомендовано для решения задач сокращения размерности данных.

Перспективы дальнейших исследований состоят в том, чтобы изучить работоспособность предложенных методов на более широком классе задач, рассмотреть применимость предложенного метода для задач с вещественным выходом.

БЛАГОДАРНОСТИ

Работа выполнена в рамках госбюджетной научно-исследовательской темы «Интеллектуальные методы и программные средства диагностирования и неразрушающего контроля качества техники военного и гражданского назначения» (гос. рег. № 0119U100360) Национального университета «Запорожская политехника» при частичной поддержке международных проектов «Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc degrees» программы «Эразмус+» Европейского Союза

и «Virtual Master Cooperation Data Science» (VIMACS)
Немецкой службы академических обменов DAAD.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Jensen R. Computational intelligence and feature selection: rough and fuzzy approaches / R. Jensen, Q. Shen. – Hoboken : John Wiley & Sons, 2008. – 300 p.
2. Subbotin S. The Dimensionality Reduction Methods Based on Computational Intelligence in Problems of Object Classification and Diagnosis / S. Subbotin, A. Oliinyk // *Recent Advances in Systems, Control and Information Technology* / Eds.: R. Szewczyk, M. Kaliczyńska. – Cham : Springer, 2017. – P. 11–19. DOI: 10.1007/978-3-319-48923-0_2
3. Subbotin S. The instance and feature selection for neural network based diagnosis of chronic obstructive bronchitis / S. Subbotin // *Applications of Computational Intelligence in Biomedical Technology*. – Cham : Springer, 2016. – P. 215–228. DOI: 10.1007/978-3-319-19147-8_13
4. Chaudhuri A. Survey sampling theory and methods / A. Chaudhuri, H. Stenger. – New York : Chapman & Hall, 2005. – 416 p. DOI: 10.1201/9781420028638
5. Subbotin S.A. Methods of sampling based on exhaustive and evolutionary search / S. A. Subbotin // *Automatic Control and Computer Sciences*. – 2013. – Vol. 47, No. 3. – P. 113–121. DOI: 10.3103/s0146411613030073
6. Lavrakas P.J. Encyclopedia of survey research methods / P. J. Lavrakas. – Thousand Oaks : Sage Publications, 2008. – Vol. 1–2. – 968 p. DOI: 10.4135/9781412963947.n159
7. Subbotin S.A. The sample properties evaluation for pattern recognition and intelligent diagnosis / S. A. Subbotin // *Digital Technologies : 10th International Conference, Zilina, 9–11 July 2014 : proceedings*. – Los Alamitos: IEEE, 2014. – P. 332–343. DOI: 10.1109/dt.2014.6868734
8. Łukasik S. An algorithm for sample and data dimensionality reduction using fast simulated annealing / S. Łukasik, P. Kulczycki // *Advanced Data Mining and Applications, Lecture Notes in Computer Science*. – Berlin : Springer, 2011. – Vol. 7120. – P. 152–161. DOI: 10.1007/978-3-642-25853-4_12
9. Subbotin S. The Sample and Instance Selection for Data Dimensionality Reduction / S. Subbotin, A. Oliinyk // *Recent Advances in Systems, Control and Information Technology*. / Eds.: R. Szewczyk, M. Kaliczyńska. – Cham : Springer, 2017. – P. 97–103. DOI: 10.1007/978-3-319-48923-0_13
10. Elavarasan N. A Survey on Feature Extraction Techniques / N. Elavarasan, K. Mani // *International Journal of Innovative Research in Computer and Communication Engineering*. – 2015. – Vol. 3, Issue 1. – P. 52–55. DOI: 10.15680/ijirccce.2015.0301009_52
11. Alpaydin E. Introduction to Machine Learning / E. Alpaydin. – London : MIT Press, 2014. – 640 p.
12. Feature Hashing for Large Scale Multitask Learning / [K. Weinberger, A. Dasgupta, J. Langford, et al.] // *26th Annual International Conference on Machine Learning (ICML '09) Montreal, June 2009 : proceedings*. – New York : ACM, 2009. – P. 1113–1120. DOI: 10.1145/1553374.1553516
13. Wolfson H. J. Geometric Hashing: An Overview / H. J. Wolfson, I. Rigoutsos // *IEEE Computational Science and Engineering*. – 1997. – Vol. 4, № 4. – P. 10–21.
14. Fast supervised discrete hashing / [J. Gui, T. Liu, Z. Sunet al.] // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2017. – Vol. 40, № 2. – P. 490–496. DOI: 10.1109/TPAMI.2017.2678475
15. Indyk P. Approximate nearest neighbors: towards removing the curse of dimensionality / P. Indyk; R. Motwani // *The 30th annual ACM symposium on Theory of computing (STOC'98), Dallas, 23–26 of May 1998 : proceedings*. – 1998. – P. 604–613. DOI:10.1145/276698.276876
16. Zhao K. Locality Preserving Hashing / K. Zhao, H. Lu, J. Mei // *Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI'14), Québec, 27–31 of July 2014 : proceedings*. – Palo Alto: AAAI Press, 2014. – P. 2874–2880.
17. Tsai Y.-H. Locality preserving hashing / Y.-H. Tsai, M.-H. Yang // *2014 IEEE International Conference on Image Processing (ICIP), Paris, 27–30 of October 2014: proceedings*. – Los Alamitos: IEEE, 2014. – P. 2988–2992. DOI: 10.1109/ICIP.2014.7025604.
18. Faure A. Perception et reconnaissance des formes / A. Faure. – Paris : Editests, 1985. – 286 p.
19. Fisher Iris dataset [Electronic resource]. – Access mode: <https://archive.ics.uci.edu/ml/datasets/Iris>
20. The plant recognition on remote sensing results by the feed-forward neural networks / [V. Dubrovin, S. Subbotin, S. Morshchavka, D. Piza] // *International Journal of Smart Engineering System Design*. – 2001. – Vol. 3, No. 4. – P. 251–256.
21. Субботин С. А. Автоматическая система обнаружения и распознавания автотранспортных средств на изображении / С. А. Субботин // *Программные продукты и системы*. – 2010. – № 1. – С. 114–116.

Received 17.07.2020.
Accepted 22.09.2020.

УДК 004.93

ОЦІНКА ІНФОРМАТИВНОСТІ І ВІДБІР ЕКЗЕМПЛЯРІВ НА ОСНОВІ ХЕШУВАННЯ

Субботин С. О. – д-р техн. наук, професор, завідувач кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна.

АНОТАЦІЯ

Актуальність. Для скорочення розмірності даних при побудові діагностичних і розпізнавальних моделей виникає необхідність відбору найбільш інформативних екземплярів, а також відбору найбільш інформативних ознак. Витрати часу на окрему реалізацію даних процедур є високими внаслідок ітеративності і взаємопов'язаності цих процедур.

Мета. Мета роботи – скорочення витрат часу на скорочення розмірності даних шляхом створення методу відбору найбільш інформативних екземплярів на основі хешування.

Метод. Запропоновано метод розрахунку ваг для визначення хешів екземплярів, який детермінованим способом визначає ваги ознак на основі їх рангів, які, у свою чергу, визначає з урахуванням кількості рівних розбиттів діапазонів ознак, мінімально достатньої для виділення кластерів на вісі ознаки з прийнятною точністю. Це дозволяє виключити необхідність ітеративного перебору різних комбінацій ознак, визначення випадкових проєкцій ознак, а також вирішення ітеративних оптимізаційних задач пошуку найкращої проєкції ознак, що істотно скорочує витрати часу на розрахунок ваг, при цьому забезпечуючи локальну чутливість хеша. Отримані хеші можливо використовувати як для відбору екземплярів, так і для відбору ознак.

Запропоновано метод визначення індивідуальної та групової значимості екземплярів вибірки, що використовує як міру подібності відстань між хешами зразків і за аналогією з методом потенціалів знаходить потенціали, що наводяться класами на кожен екземпляр, а на їх основі визначає показники значущості екземплярів, виходячи з того, що екземпляр в просторі ознак тим інформативніше, чим менше мінімальна різниця потенціалів класів, що наводяться на екземпляр.

Запропоновано метод визначення оцінок інформативності ознак, який на основі нормування ваг, отриманих при формуванні хешів, визначає показники інформативності ознак, віддаючи перевагу ознаками з меншою кількістю розбиттів.

Результати. Проведено експериментальне дослідження, яке підтвердило працездатність запропонованих методів при вирішенні практичних завдань.

Висновки. Розроблене математичне забезпечення може бути рекомендовано для вирішення завдань скорочення розмірності даних.

КЛЮЧОВІ СЛОВА: екземпляр, ознака, інформативність, хешування, хеш, скорочення розмірності вибірки.

UDC 004.93

EVALUATION OF INFORMATIVITY AND SELECTION OF INSTANCES BASED ON HASHING

Subbotin S. A. – Dr. Sc., Professor, Head of the Department of Software Tools at the National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

ABSTRACT

Context. To reduce the data dimensionality in the diagnostic and recognition model construction, it becomes necessary to select the most informative instances, as well as to select the most informative features. The time spent on the separate implementation of these procedures is high due to the iterativity and interconnectedness of these procedures.

Objective. The purpose of this work is to reduce the time spent on reducing the data dimensionality by creating a method for selecting the most informative instances based on hashing.

Method. A method for calculating weights for determining the hashes of instances is proposed, which determines the weights of features based on their ranks in a deterministic way, which, in turn, determines, taking into account the number of equal partitions of the ranges of features, the minimum sufficient to distinguish clusters on the axis of the feature with acceptable accuracy. This eliminates the need for iterative enumeration of various combinations of features, determining random projections of features, as well as solving iterative optimization problems of finding the best projection of features, which significantly reduces the time spent on calculating weights, while ensuring the local sensitivity of the hash. The hashes obtained can be used both for the selection of instances and for the selection of features.

A method for determining the individual and group significance of sample instances is proposed, in which it uses the distance between the hashes of the instances as a measure of similarity and, by analogy with the potential method, finds the potentials induced by the classes for each instance, and on their basis determines the indicators of the significance of the instances, based on the fact that the instance in the feature space, the more informative the less the minimum potential difference of the classes induced on the specimen.

A method for determining the estimates of the informativeness of features is proposed, which, on the basis of normalizing the weights obtained during the formation of hashes, determines the indicators of the informativeness of features, giving preference to features with a smaller number of partitions.

Results. An experimental study has been carried out, which has confirmed the efficiency of the proposed methods in solving practical problems.

Conclusions. The developed software can be recommended for solving problems of data dimension reduction.

KEYWORDS: instance, attribute, informativeness, hashing, hash, reduction of the sample size.

REFERENCES

1. Jensen R., Shen Q. Computational intelligence and feature selection: rough and fuzzy approaches. Hoboken, John Wiley & Sons, 2008, 300 p.
2. Subbotin S., Oliinyk A. Eds.: Szewczyk R., Kaliczyńska M. The Dimensionality Reduction Methods Based on Computational Intelligence in Problems of Object Classification and Diagnosis, *Recent Advances in Systems, Control and Information Technology*. Cham, Springer, 2017, pp. 11–19. DOI: 10.1007/978-3-319-48923-0_2
3. Subbotin S. The instance and feature selection for neural network based diagnosis of chronic obstructive bronchi-

- tis, *Applications of Computational Intelligence in Biomedical Technology*. Cham, Springer, 2016, pp. 215–228. DOI: 10.1007/978-3-319-19147-8_13
4. Chaudhuri A., Stenger H. Survey sampling theory and methods. New York, Chapman & Hall, 2005, 416 p. DOI: 10.1201/9781420028638
 5. Subbotin S.A. Methods of sampling based on exhaustive and evolutionary search, *Automatic Control and Computer Sciences*, 2013, Vol. 47, No. 3, pp. 113–121. DOI: 10.3103/s0146411613030073
 6. Lavrakas P.J. Encyclopedia of survey research methods. Thousand Oaks, Sage Publications, 2008, Vol. 1–2, 968 p. DOI: 10.4135/9781412963947.n159
 7. Subbotin S.A. The sample properties evaluation for pattern recognition and intelligent diagnosis, *Digital Technologies : 10th International Conference, Zilina, 9–11 July 2014 : proceedings*. Los Alamitos, IEEE, 2014, pp. 332–343. DOI: 10.1109/dt.2014.6868734
 8. Lukasik S., Kulczycki P. An algorithm for sample and data dimensionality reduction using fast simulated annealing, *Advanced Data Mining and Applications, Lecture Notes in Computer Science*. Berlin, Springer, 2011, Vol. 7120, pp. 152–161. DOI: 10.1007/978-3-642-25853-4_12
 9. Subbotin S., Oliinyk A. Eds.: R. Szewczyk, M. Kaliczyńska The Sample and Instance Selection for Data Dimensionality Reduction, *Recent Advances in Systems, Control and Information Technology*. Cham, Springer, 2017, pp. 97–103. DOI: 10.1007/978-3-319-48923-0_13
 10. Elavarasan N., Mani K. A Survey on Feature Extraction Techniques, *International Journal of Innovative Research in Computer and Communication Engineering*, 2015, Vol. 3, Issue 1, pp. 52–55. DOI: 10.15680/ijirce.2015.0301009_52
 11. Alpaydin E. Introduction to Machine Learning. London, MIT Press, 2014, 640 p.
 12. Weinberger K., Dasgupta A., Langford J., Smola A., Attenberg J. Feature Hashing for Large Scale Multitask Learning, *26th Annual International Conference on Machine Learning (ICML '09) Montreal, June 2009 : proceedings*. New York: ACM, 2009, pp. 1113–1120. DOI: 10.1145/1553374.1553516
 13. Wolfson H. J., Rigoutsos I. Geometric Hashing: An Overview, *IEEE Computational Science and Engineering*, 1997, Vol. 4, № 4, pp. 10–21.
 14. Gui J., Liu T., Sun Z., Tao D., Tan T. Fast supervised discrete hashing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, Vol. 40, No. 2, pp. 490–496. DOI: 10.1109/TPAMI.2017.2678475
 15. Indyk P., Motwani R. Approximate nearest neighbors: towards removing the curse of dimensionality, *The 30th annual ACM symposium on Theory of computing (STOC'98), Dallas, 23–26 of May 1998 : proceedings. – 1998*, pp. 604–613. DOI:10.1145/276698.276876
 16. Zhao K., Lu H., Mei J. Locality Preserving Hashing, *Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI'14), Québec, 27–31 of July 2014 : proceedings*. Palo Alto, AAAI Press, 2014, pp. 2874–2880.
 17. Tsai Y.-H., Yang M.-H. Locality preserving hashing, *2014 IEEE International Conference on Image Processing (ICIP), Paris, 27–30 of October 2014: proceedings*. Los Alamitos, IEEE, 2014, pp. 2988–2992. DOI: 10.1109/ICIP.2014.7025604.
 18. Faure A. Perception et reconnaissance des formes. Paris, Editests, 1985, 286 p.
 19. Fisher Iris dataset [Electronic resource]. Access mode: <https://archive.ics.uci.edu/ml/datasets/Iris>
 20. Dubrovin V., Subbotin S., Morshchavka S., Piza D. The plant recognition on remote sensing results by the feed-forward neural networks, *International Journal of Smart Engineering System Design*, 2001, Vol. 3, No. 4, pp. 251–256.
 21. Subbotin S. A. Avtomaticheskaja sistema obnaruzhenija i raspoznavanija avtotransportnyh sredstv na izobrazhenii, *Programmnye produkty i sistemy*, 2010, No. 1, pp. 114–116.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

PROGRESSIVE INFORMATION TECHNOLOGIES

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

UDC 51-74:004.021

SOLUTION OF THE PROBLEM OF PLACING MEDICAL FACILITIES IN CITY DEVELOPMENT PROJECTS

Danshyna S. Yu. – PhD, Associate Professor, Associate Professor of the Department of geoinformation technologies and space monitoring of the Earth, National Aerospace University named after N. Zhukovsky “Kharkiv Aviation Institute”, Kharkiv, Ukraine.

Nechausov A. S. – PhD, Associate Professor of the Department of geoinformation technologies and space monitoring of the Earth, National Aerospace University named after N. Zhukovsky “Kharkiv Aviation Institute”, Kharkiv, Ukraine.

ABSTRACT

Context. The problem of rational allocation of medical facilities of inhabited locality was considered. Methods of evaluating the effectiveness of using the existing medical network and finding ways to improve it when implementing urban development projects were proposed.

Objective. The goal of the work is to build and study the procedure for solving the problem of placing medical facilities, considering the existing infrastructure of the city to fulfill the accessibility requirements.

Method. A set of factors that affect the placement of a medical facilities and allow a systematic and reasonable decision to be made in choosing a location has been formed. A method for solving the problem of the location of medical facilities, providing an increase in their accessibility level considering the population, its spatial distribution, existing in the settlement of road junction and traffic congestion is proposed. The procedure for solving the problem is presented in the form of an IDEF0 model. The method is based on geoinformation analysis of data, the results of which by spatial clustering are presented in the form of a set of cartographic models, the aggregate of which allows to form a decision about the location. The method allows to assess the accessibility level of existing medical facilities in the inhabited locality, create a list of places for their possible location, find areas of the locality that are not in the access zone.

Results. The method of choosing the location of the medical center has been improved, which makes it possible to make decisions based not only on the shortest distance from the center to the patient, but also on the level of its accessibility. For the first time, the structure of information technology for multivariate analysis of a network of city medical centers for decision support systems using GIS technologies is proposed.

Conclusions. The medical facilities location problem, which is in increasing the level of accessibility by expanding the area of medical coverage of the territory in which the city population lives by using GIS was practically fulfilled as an IDEF0 model that defines the procedure for solving the problem and choosing the location of the medical center. The practical significance of obtained results has been proven on a practical example – an analysis of the existing hospital system for the provision of emergency medical care in Kharkiv.

KEYWORDS: accessibility level, coverage area, population, patient delivery time, IDEF0 model, geoinformation technologies, network analysis, ArcGIS.

ABBREVIATIONS

GIS is a geoinformation system.

NOMENCLATURE

a – rate of housing area per person, m^2 /people;

β – statistical distance influence coefficient;

d_k – distance to the k -th patient;

$\{E_k\}$ – patient of the city plurality;

F – volume of housing;

G – target function depending on the distance from the medical center to the location of patients;

N – number of patients requiring the services of a medical institution (population of the microdistrict);

P_k – number of services that a medical institution provides for all k -patients which need services;

R – accessibility level;

S_n – area in which patients live;

S_o – medical coverage area;

S – microdistrict area, ha;

S_{norm} – normative housing density, considering the number of storeys, depending on the construction of the city zoning, m²/ha;

t – time, the value of which depends on the maximum permissible distance between the patient in need of the service and the medical center providing the service;

$t^{-\beta}$ – rate of time to cover the distance between the patient in need of the service and the medical center providing the service;

(x, y) – medical center coordinates;

(x_k, y_k) – patient location coordinates, characterizing the structure of their settlement in the city;

Z – the maximum number of patients who can be simultaneously provided with services in a medical center.

INTRODUCTION

Urban development is a natural process. City development projects are aimed at improving the industrial, economic, social sphere. However, in recent decades, the social sphere of Ukrainian cities has developed and functioned very unevenly, leading to a significant imbalance in the range of services provided to the population. This unevenness is manifested in the massive development of trade, recreation and entertainment centers, fast food enterprises. At the same time, there is stagnation and destruction in housing and communal services, consumer services, sports and medical fields, etc. The following emphasizes the particular importance and relevance of problems in this area [1], [2]:

– financing according to the “residual principle”, the lack of opportunities for decent maintenance and repair leads to the fact that many social facilities are liquidated, their purpose is changed, depreciation of fixed assets increases, the quality of services provided decreases;

– the pace of development of social infrastructure is significantly lower than the pace of socio-economic and cultural changes, which means that the existing infrastructure is not able to provide favorable living conditions, improve the quality of life, etc.;

– one of the most important characteristics of social infrastructure is the territorial distribution, which is affected not only by local social policy but also by regional and state social policy;

– the growing need of social services requires the development and implementation of national policies aimed at the development of social infrastructure, which should reflect the specifics of each city, implementing specific activities in each region and city;

– irrational spatial distribution of objects of the social sphere complicates their use by the population, which is significantly aggravated by the suboptimal work of public transport, violating the principle of “walking accessibility”, wasting of free time, etc.

The healthcare sector deserves special attention, which is a qualitative characteristic of the economically active population: the state of the nation’s health directly affects

labor productivity and the prospects for socio-economic development of society [3]. However, according to sociologists, 76.4% of those who took part in the survey see problems in the field of medical care, and every fifth considers problems in the health care system as personal [4]. All this is compounded by the ongoing medical reform in the country, the results of which, to date, have received mixed reviews, both from specialists and from the government and the population.

These reasons lead to attention level growth to the problem of finding effective ways to use the existing medical infrastructure and finding directions for its enhancement to further improve the living conditions of the population within the urban development projects.

The object of study in this work are models for evaluating the effectiveness of using the existing medical network and finding ways to improve it when implementing urban development projects.

The subject of study is the model of placing a medical facility task.

The purpose of the work is to build and study the procedure for solving the problem of placing medical facilities, considering the existing infrastructure of the city to fulfill the accessibility requirement. To achieve this goal, it is necessary to solve the following tasks:

– consider the factors affecting the choice of a medical institution, choose the most important of them;

– formulate the problem of the placement of medical facilities, considering the selected factors;

– develop a procedure for solving the problem of placement;

– evaluate the possibility of applying the procedure for solving the problem of placement in urban development projects;

– are models for evaluating the effectiveness of using the existing medical network and finding ways to improve it when implementing urban development projects.

1 PROBLEM STATEMENT

A network of health care institutions is a set of city institutions that satisfy the population’s need for medical care as a result of the professional activities of medical workers. The functioning of such a network is performed considering the requirements of the population in medical care, its timeliness and accessibility.

Accessibility level – the ability to provide patients with the necessary range of services of a certain quality with minimal time spent.

Suppose:

– all residents living in the territory under consideration are potential patients of a medical institution and form plurality of patients $\{E_k\}$ ($k = \overline{1, N}$);

– the specifics of medical services are not considered, it is believed that the medical center is able to provide any medical service that the patient needs, but the total number of services is limited;

– there is a “limit of possibilities” for a medical center, the value of which depends on the maximum number of

patients who can be provided with medical services at the same time;

- the quality of medical services, the reputation of the doctor, etc., are not considered, i.e. the decision to choose a specific medical center is affected only by its proximity to the patient;

- there is no competition for patients within the network.

The network elements – medical centers – are placed in such a way that the value of the objective function, which depends on the distances from the medical center to the location of the patients, is minimal, i.e.

$$G = \sum_{k=1}^N d_k \rightarrow \min . \quad (1)$$

Given that the level of accessibility will be maximum, i.e.

$$R = \frac{\sum_{k \in E_k} P_k}{N} \cdot 100\% \rightarrow \max , \text{ at } \sum_{k \in E_k} P_k = Z . \quad (2)$$

Under given conditions, it is necessary to obtain the coordinates of the medical center.

2 REVIEW OF THE LITERATURE

The strategic goal of state policy in the field of health is to improve the health stat of the population, increase the real availability of medical care for all its layers. Development projects of many cities are aimed at structural, organizational and functional restructuring of the medical care system by creating and/or improving the network of healthcare institutions [3]. In this regard, the goal of development projects is to increase the availability of medical center that provide diverse medical care to the urban population.

In modern literature, the accessibility of a medical center is interpreted differently [3], [5], [6], which emphasizes the multifactorial nature of this concept, its dependence on objective, subjective, external and internal factors. In conditions of equal access, considering accessibility as an opportunity to provide the necessary range of services of a certain quality with minimal time and money [3], [5], the factor of placement of a medical center as a geographical object comes out on top, emphasizing the relationship between location and medical services choice [7], [8].

Problem (1) is known in the literature as the P-median model described in 1964 by S. Hakimi, who suggested that social institutions should be located so that the sum of all the distances from demand points to point P (social service supply points) should be minimal [2].

Given that the medical center is a geographical object that has certain coordinates (x, y) , we will represent the plurality of patients of the city $\{E_k\}$ as points of Cartesian space and place them on the plane (Fig. 1).

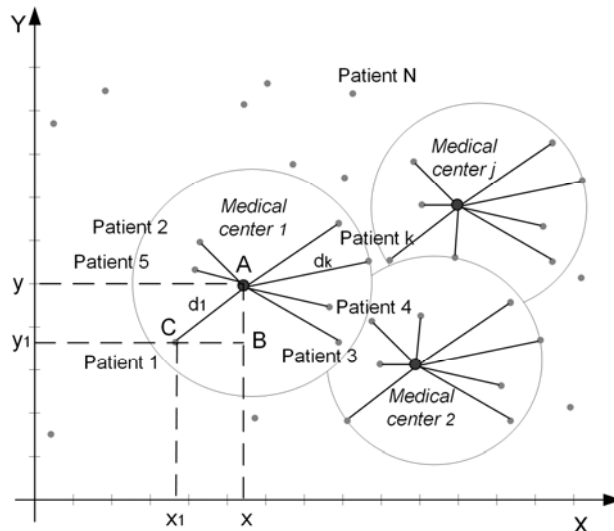


Figure 1 – Geometric interpretation of the medical center placing problem

Then the distance d_k is the distance from the center to the k -th patient, the value of which can be found as the length of the segment on the coordinate plane from the known coordinates of its ends. For example, we find the length of the segment d_1 from a right triangle ΔABC as the length of the hypotenuse:

$$d_1 = AC = \sqrt{(x - x_1)^2 + (y - y_1)^2} . \quad (3)$$

Substituting the expression (3) in the formula (1), the objective function will be represented in the form:

$$G = \sum_{k=1}^N \sqrt{(x - x_k)^2 + (y - y_k)^2} \rightarrow \min . \quad (4)$$

Wherein, it can be assumed that the set of patient coordinates (x_k, y_k) ($k = \overline{1, N}$) determines the structure of their settlement, and the obtained coordinates of the medical center form the structure of the network of medical centers in the city.

Among the main disadvantages of the multicriteria optimization task (4) the following can be noted [2], [8]:

- the high dimensionality of the problem, due to the fact that the set $\{E_k\}$ is finite, but its capacity can reach tens of thousands and even millions of units;
- in conditions of urban settlement (possibly uneven), it is difficult to consider the coordinates of all patients;
- in the statement of the problem (3) there is no concept of the “limit of possibilities” of the medical center, which leads to a unilateral consideration of the problem;
- in reality (on the ground) the “shortest” roads may be absent.

The latter was confirmed by field studies [2], [8]: the real zones of the medical services provision point location significantly differed from the results obtained using the model (4). And although it was proposed to adjust the

results using coefficients that consider transport, social, political and other factors [6], [8], but this did not help to solve the remaining problems.

A fundamentally different approach to solving the problem of the placement of social objects, focusing on spatial accessibility, was proposed by Yue Zhang et al. [9]. The authors prove that the choice of the location of the medical center is primarily affected by the population, and the level of accessibility (2), in this case, it should be regarded as the ratio of supply to demand.

Research of I. N. Sener et al. [10] show that the closer the medical center to the patient is, the more accessible it is for him. On the other hand, it was proved in [11] that the more patients assigned to this center, the less accessible it is for them. Therefore, expression (2) can be represented as follows:

$$R = \frac{\sum_{k \in E_k} P_k t_k^{-\beta}}{Nt} \cdot 100\% . \quad (5)$$

The choice of the medical center location in (5) is made from the stipulation of maximizing the demand for its services, considering the general limit of the possibilities for providing medical services. At the same time, the travel time determined by the direct relative position of the patient and the medical center is the main factor by which the place of receiving medical services is selected [9], [11]. Given the existing settlement structure and population, this factor has significant weight. However, its significance can be offset by other factors. Among such factors, one can consider the presence of communication lines and their congestion, the nature of settlement, the presence of personal transport, the available public transport routes, etc.

The problem of spatial accessibility cannot be solved within the framework of the healthcare system alone. It requires improving the quality and standards of the population living, improving and modernizing the transport infrastructure as a connecting link [6], [12]. Focusing on the transport infrastructure providing equal access to the services of the medical center, we assume that the value of the coefficient β is determined by its parameters, in particular, the existing roads and their congestion.

Analyzing fig. 1, the numerator of function (5) can be interpreted as the number of services the medical center can provide for patients living within a circle with a radius, the value of which depends on time $t^{-\beta}$. In this case, the denominator can be considered as the number of patients living within a circle with a radius, the value of which depends on the time for which the patient will overcome the maximum permissible distance from the place of residence to the medical center;

Among the main disadvantages of problem (5), we note the following:

- an indefinite algorithm for evaluating $t^{-\beta}$ and t ;
- lack of explanation in determining the coefficient β ;

– in the statement of the problem there is no notion of the medical center coordinates, which makes it difficult to solve the problem of placement.

Nevertheless, the analysis of solving the problem approaches allows you to create a set of factors affecting the placement and choice of a medical institution. These include:

- the population (as the potential number of patients) living within the territory, the size of which is determined by the maximum travel time to the medical center and the structure of their settlement;
- the number of services that a medical center can provide for patients in the territory in question;
- the norm of time to overcome the distance between the patient in need of the service and the medical center providing the service;
- really existing road junction, roads and their congestion.

These factors must be considered when constructing the procedure for solving the problem of the placement of medical institutions.

3 MATERIALS AND METHODS

In accordance with State Building Codes of Ukraine B.2.2-12:2018 “Planning and development of territories”, the population of a city district is determined depending on the housing stock and housing standards per inhabitant and is calculated according to the formula

$$N = \frac{F}{a} = \frac{S \cdot S_{norm}}{a} . \quad (6)$$

In this case, the demand for the services of a medical center is determined by the number of people (patients) living in a certain territory. Denote the area of this territory as S_n . On the other hand, the number of services a medical center provides for patients residing within a certain territory is determined by the distance that a patient travel from home to the center for a given period of time. The area, the value of which depends on this distance, we will consider as the area covered by medical services and denote S_o .

Thus, expression (5) can be represented as follows:

$$R = \frac{S_o}{S_n} \cdot 100\% . \quad (7)$$

In this case, task (7) considers the maximum number of factors affecting the placement and choice of a medical institution.

The process of solving this problem is presented in the form of a context diagram (Fig. 2), created on the basis of the IDEF 0 standard, which shows a generalized process model, defines a single point of view, subject and purpose of modeling in accordance with the requirements of current standards and norms.

The attributes of the generalized model are:

1. Input data:
 - Map of the city;

- Location of the city microdistricts;
- Coordinates of existing medical centers of the city;
- Congestion of city roads;
- City road map or transport infrastructure;

2. Control:

- Norm of living space per person;
- Normative building density;
- Requirements for the location of the medical center determined by the current legislation in the field of healthcare;
- Time standards that determine the necessary speed of patient delivery to medical center;
- Comparison criteria.

3. Output:

- Evaluation of the availability of existing medical centers;
- Recommended placements of medical center;
- The list of city areas that did not fall into availability zone.

It is easy to notice that the attributes of the model (Fig. 2) contain spatially distributed data, the effective processing of which is possible by using GIS.

GIS are information systems that allow you to collect, organize, analyze, distribute spatial information to make the necessary decisions. Spatial or GIS analysis of multicriteria decisions includes a set of methods and tools for combining geospatial data with the arguments of decision makers [13], [14]. Given the complex nature of spatial solutions, we single out a number of functional areas through which the output data set is formed. We will form an IDEF0 model of the procedure for solving the problem (7) and choosing the location of the medical center using GIS (Fig. 3). Note that for the convenience of further analysis and decision making the given IDEF0-model is provided with a representation of intermediate results in the form of cartographic models, obtained using the spatial clustering method. This method has received positive feedback among experts and has proven itself in solving environmental, social and political problems [12], [13].

The implementation of proposed procedure is possible by using the ArcGIS software suite by Esri.

ArcGIS is an effective universal solution for simultaneous work with the geospatial component, multimodal network datasets, and powerful tools of analytics, which helps to perform not only the fundamental spatial analysis, but also to obtain spatio-temporal models on a cartographic basis [12], [15].

When performing the steps, described in the IDEF0 model the ArcGIS Network Analyst extension was used, which helps to answer questions like the following:

- What is the quickest way to get from point A to point B?
- Which houses are within five minutes of a medical center?
- What are live traffic conditions like, and how do they affect network analysis results?

Public services, municipal and other organizations benefit from the ArcGIS Network Analyst extension be-

cause it helps them run their operations more efficiently and make better strategic decisions [13], [14].

In ArcGIS a network service area is a region that encompasses all accessible streets, that is, streets that lie within a specified impedance. For instance, the 10-minute service area for a facility includes all the streets that can be reached within 10 minutes from that facility. The accessibility parameter refers to how easy it is to go to a site.

In Network Analyst, accessibility can be measured in terms of travel time, distance, or any other impedance on the network. With Network Analyst, service areas around any location on a network can be found. A network service area is a region that encompasses all accessible streets, that is, streets that lie within a specified impedance. For instance, the 10-minute service area for a facility includes all the streets that can be reached within 10 minutes from that facility.

One simple way to evaluate accessibility is by a buffer distance around a point. However, considering people travel by road, this method won't reflect the actual accessibility to the site. Service networks computed by Network Analyst can overcome this limitation by identifying the accessible streets within five kilometers of a site via the road network. Once created, service networks can be used to see what is alongside the accessible streets.

4 EXPERIMENTS

The possibility of using the proposed model (7) in projects for the development of social infrastructure in Kharkiv was considered. The model was studied while solving the problem of analyzing the existing system of providing emergency care in Kharkiv. The following were used as initial data:

- Open Street Map basemap layers;
- Addresses and names of Kharkiv emergency medical institutions;
- Data on the city's road transport network (traffic congestion, road signs, lane width).
- Vector *.shp layers with the marked boundaries of the microdistricts of Kharkiv;
- Database of residential buildings storeys number (for calculating the population).

To verify the proposed model, we determined the population size and the area of the territory where patients live. The results obtained showed conformity with the data of the statistical office of the Kharkiv region, which made it possible to conduct an experiment to study the effectiveness of the proposed approach in real conditions and on real objects.

As a result of the experiment, the coverage areas of emergency care services provided by hospitals in Kharkiv microdistricts and nearby villages were obtained, as a result of which it became possible to assess and compare the level of accessibility of emergency hospitals. The results obtained showed a low level of logistics in the distribution of objects of the existing system of emergency care in the city.

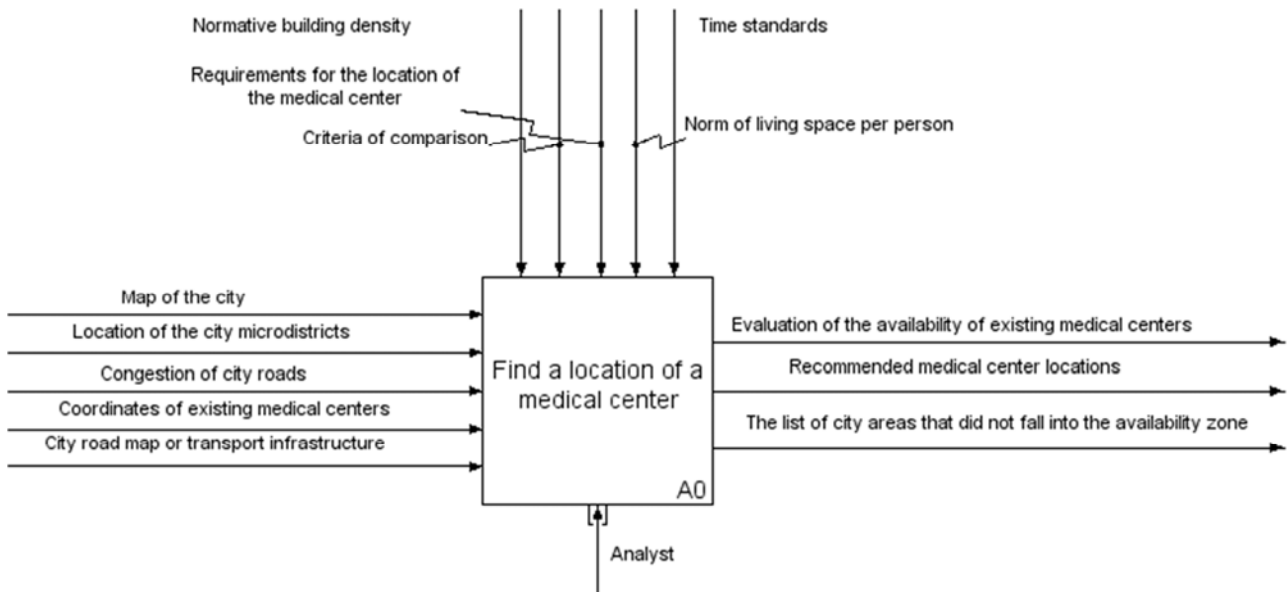


Figure 2 – Context diagram process of finding a location of a medical center

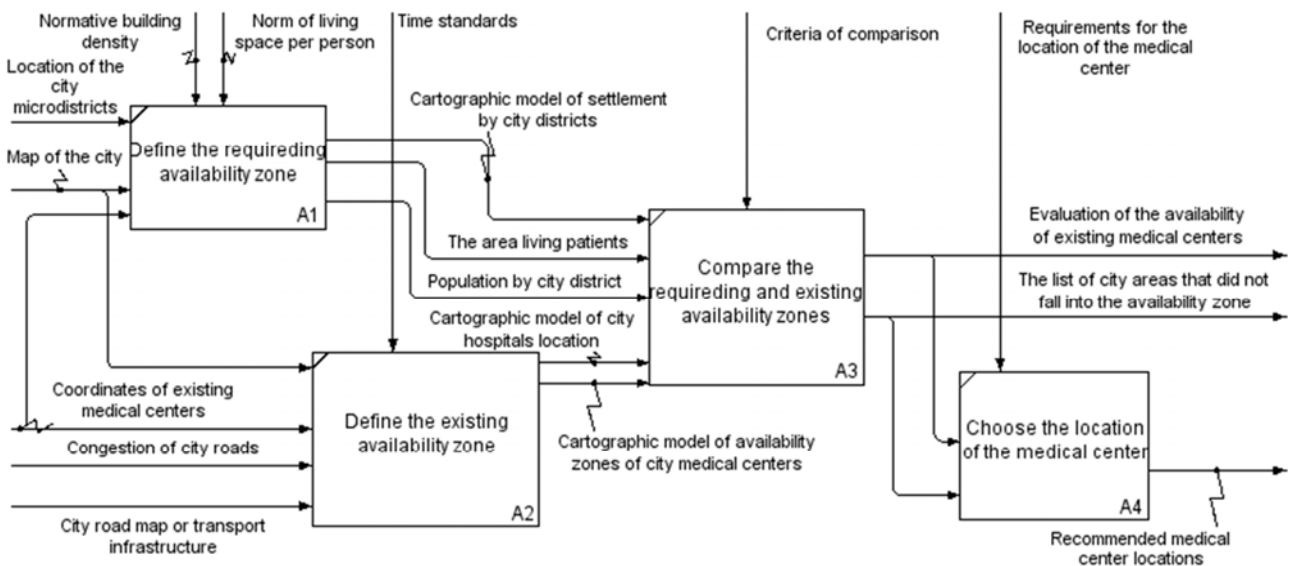


Figure 3 – IDEF0-model of the procedure for solving the problem of placing a medical center in the city

To select possible directions of development, a scenario for correcting this system by changing the location of some hospitals, considering the population size and the existing road junction, is proposed.

After the correction, assessments of the level of accessibility of emergency hospitals in Kharkiv were obtained, indicating an increase in the number of patients who will receive medical care timely.

5 RESULTS

Consider the problem of analyzing the existing system of providing emergency care in Kharkiv and choosing possible directions for its solution.

In accordance with the proposed approach, the first stage (Fig. 3) of the IDEF0-model is determination of the necessary accessibility zone of the medical center. The

implementation of this stage was held in the following sequence:

1. According to the Kharkiv City Council, Kharkiv is divided into 56 microdistricts. These microdistricts and nearby villages serve 17 hospitals for emergency and ambulance medical care. Using ArcGIS, we map the regions and towns under consideration on a map of Kharkiv. Using the Database “Ambulance Hospital Addresses”, we will show the location of the hospitals on a map of the districts and get a cartographic model of ambulance hospitals location in Kharkiv city (Fig. 4).

2. Using Google Planet, we determine the area of the microdistricts of Kharkiv and adjust them considering the population of the territory. For example, in the greater part of Pomerki microdistrict, there is a plot of forest – the botanical monument “Pomerki”. In this case, the residen-

tial area of the microdistrict is only 11% (about 13.32 hectares). As a result, the total area of the territory where patients live (S_n) amounted to 356.46 km².

By the formula (6) we find the number of inhabitants of these districts. Moreover, for districts where the number of storeys is higher than five $S_{norm} = 3000$ m²/ha, with one-, two-story buildings, $S_{norm} = 1800$ m²/ha, for the mixed development zone $S_{norm} = 2300$ m²/ha. We will form the database “Population of the districts of Kharkiv”, which shows the distribution of 1.419 million people in the districts.

3. Using ArcGIS, we will construct a cartographic model of the distribution of residents among the districts of Kharkiv based on the database “Population of the Kharkiv districts”, where polygons with a larger population are darker, and the least populated areas are lighter (Fig. 5).

At the second stage (Fig. 3) of solving the objective, we determine the accessibility zones of existing ambulance and emergency care hospitals. In this case, we will consider the following.

1. Statistical data (infoportal.ua) of traffic congestion in Kharkiv shows that during the day the situation on the roads can change from light (1 point) to heavy traffic (“traffic jam” – 4 points). An analysis of the congestion of roads leading to the considered hospitals showed that on

average the worst situation (“obstructed” – 3 points) is observed from 8:00 to 9:00, from 12:00 to 13:00, from 16:00 to 17:00. In availability assessment we will consider precisely these time intervals.

2. The existing time standards determine that the patient must be delivered to the hospital within 10 minutes. For example, in [16] it is noted that transportation for a period of more than 10–20 minutes shifts the severity of the victim up one category. So, a slightly wounded person within 20–30 minutes of transportation is likely to go into the category of victims of medium severity. Therefore, when assessing the availability, we will consider the delivery time of 5 and 10 minutes.

3. Under the assumptions made, in ArcGIS Network Analyst we will create cartographic models of hospital access zones (Fig. 6). Using the function “service areas” (availability zones) for vehicles we define the area of coverage with medical services (S_o) for the selected time intervals, considering different transportation times (table 1) [12], [17].

At the third stage (Fig. 2), we compare the obtained results and by using (6) we find the level of accessibility of emergency hospitals for the selected time period by vehicles (Table 2).

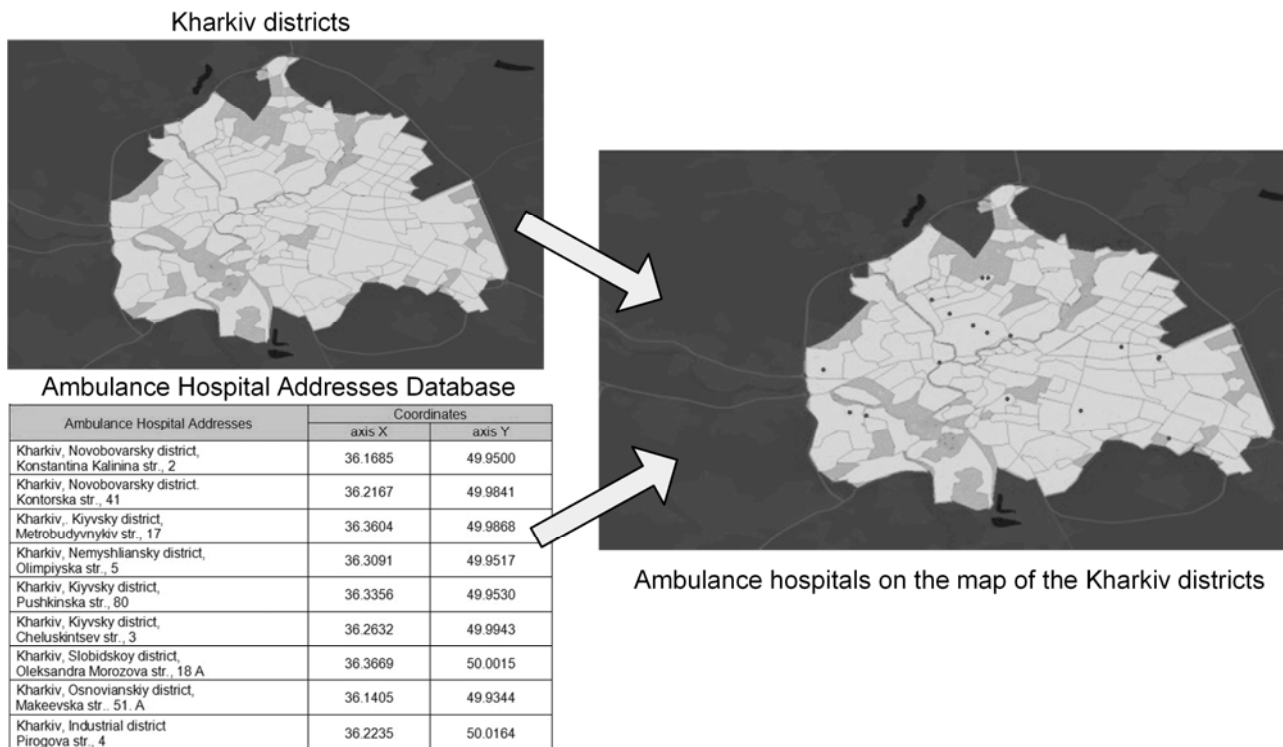


Figure 4 – Cartographic model of ambulance hospitals location in Kharkiv city

Database "Population of the Kharkiv districts "

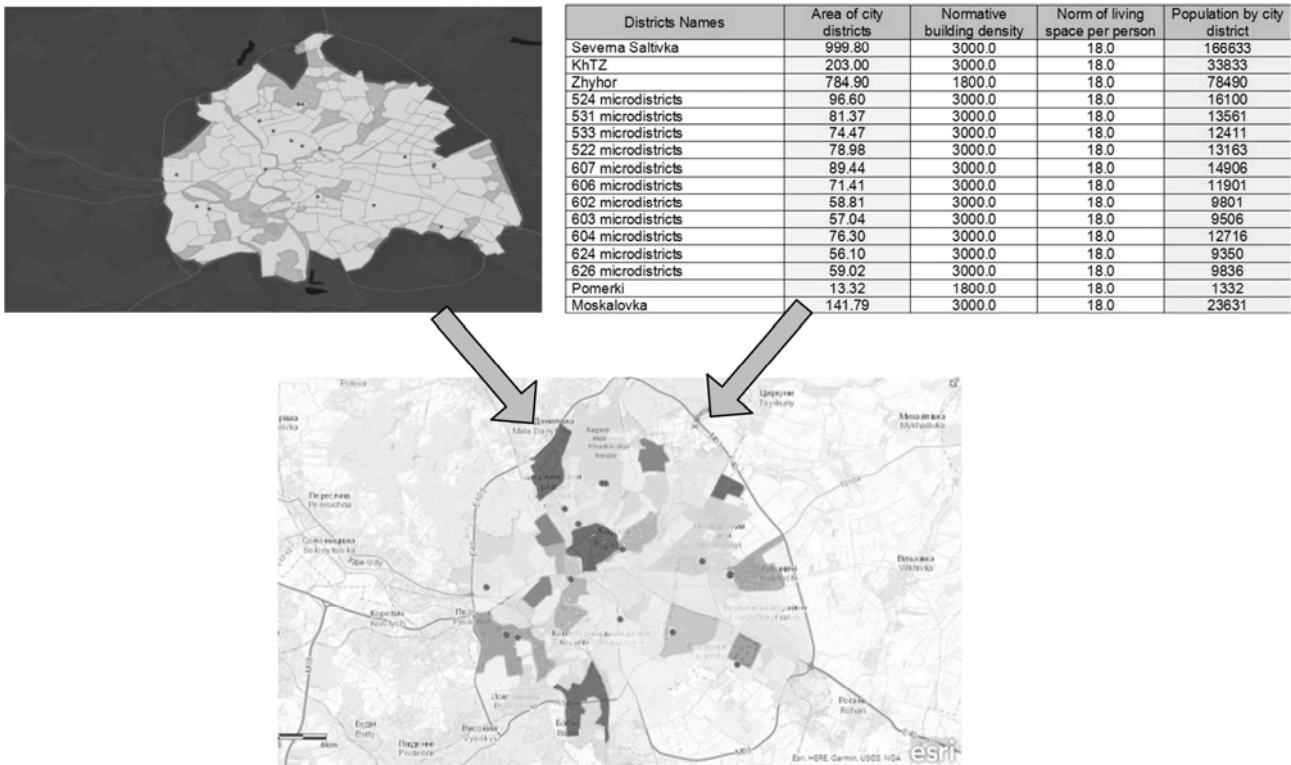


Figure 5 – Cartographic model of settlement by Kharkiv city

Table 1 – The services of emergency hospitals in the districts of Kharkov and nearby villages coverage area

Norm of patient delivery time, min	Time period		
	8-00	12-00	16-00
5	76.02 km ²	65.9 km ²	66.86 km ²
10	219.9 km ²	211.7 km ²	211.01 km ²

Table 2 – The level of accessibility of emergency hospitals in Kharkiv

Norm of patient delivery time, min	Time period		
	8-00	12-00	16-00
5	21.33 %	18.49 %	18.76 %
10	61.69 %	59.39 %	59.20 %

The data obtained (Fig. 6) make it possible to find the districts of Kharkiv and nearby villages that did not fall into the accessibility zone of the considered hospitals. Seven districts fell here: Zhyhor with a population of 78,900 people, Pyatikhatki with a population of 9,800 people, Bolshaya Danilovka (6,000 people), Bolshaya Panasovka (2,934 people), Guty (1,007 people), Novy Korotich (840 people) and Tyschenky (112 people). Figure 7 shows how, using the obtained cartographic models, for the 15-minute drivetime accessibility zone, the regions that did not fall into the accessibility zone were determined.

The results obtained are the basis for the analysis of the effectiveness of the existing medical infrastructure of the city. However, in the IDEF0 model, the fourth stage is foreseen, related to the medical center location selection process.

Based on the results, we assume that it is possible to change the location of some hospitals, for example, by organizing a feldsher-midwife station or branch of a hospital to provide emergency medical care. Considering

population and existing denouement, three hospitals “moved” closer to the ring road.

Repeating the second and third stages of the procedure for solving the placement objective (Fig. 3), new values were obtained for the coverage area and the level of accessibility of hospitals (Table 3).

6 DISCUSSION

Summarizing the results, we can draw the following conclusions. Preliminary estimates of the area of the territory where patients live (S_n) and the population (N) obtained at the first stage of the IDEF model do not contradict the data of the Main Regional Statistics Office of Kharkiv (<http://kh.ukrstat.gov.ua/>), which indicates a high reliability of the results. In particular, the error in determining the value was about 1.8% (upward), in the definition of N value – 1.7% (downward). At the same time, the used spatial clustering method depends only on distance, generates repeatable results, is free from subjectivity and does not depend on the experience of observers. That is why, its further application at the second stage of the

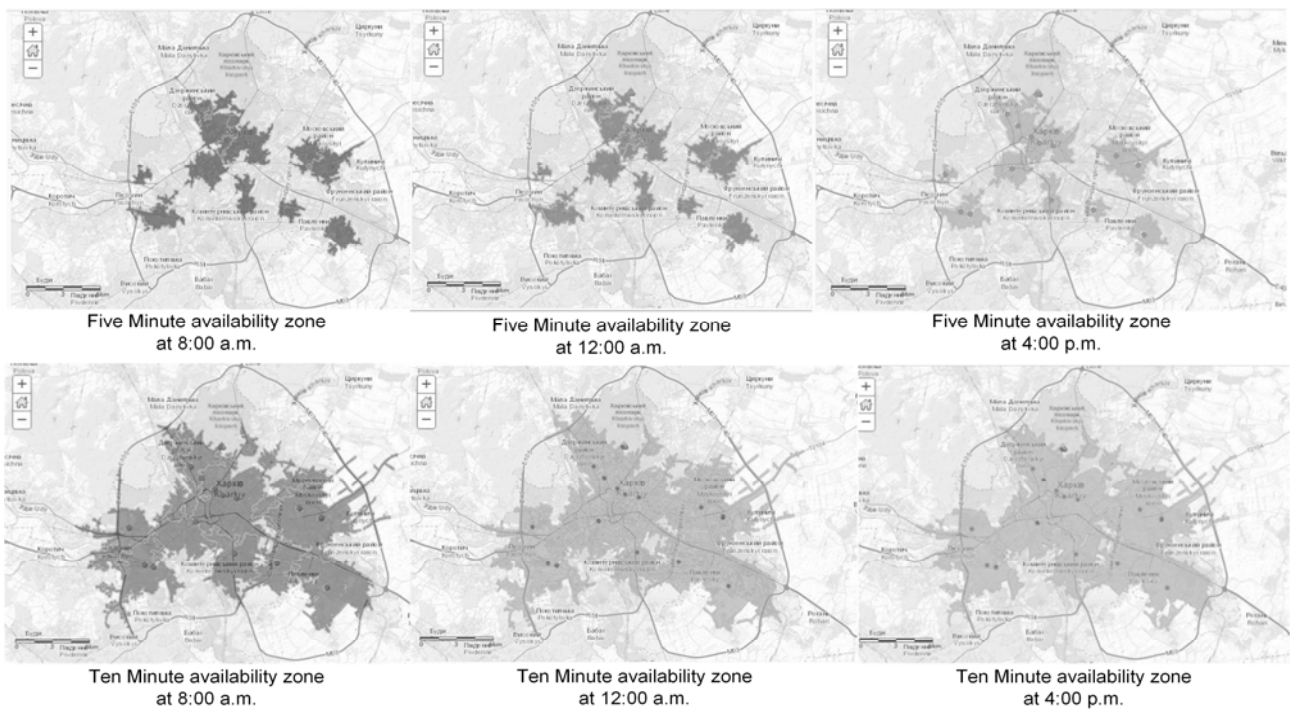


Figure 6 – Cartographic model of availability zones of ambulance hospitals by Kharkiv city

Table 3 – Coverage area and accessibility level of emergency hospitals in Kharkiv after adjusting the location of some of them

Norm of patient delivery time, min	Hospital services coverage			Hospital accessibility level		
	Time period			Time period		
	8-00	12-00	16-00	8-00	12-00	16-00
5	83.62 km ²	75.16 km ²	75.45 km ²	23.46 %	21.08 %	21.17 %
10	346.14 km ²	329.08 km ²	329.22 km ²	97.10 %	92.32 %	92.36 %

IDEF model is justified and allows us to accept the obtained values of the medical services coverage area (S_o).

We also note that the efficiency and high accuracy of the spatial clustering method, the wide use of which is proposed in the IDEF model, is confirmed by foreign studies conducted in various fields [12], [13]. Thus, the results can be considered reliable, and such that they do not contradict international experience.

The results of the study showed unevenness in the resettlement and location of existing medical centers. In particular, the most populated microdistricts are Severnaya Saltovka (166633 people), Alekseyevka (100500 people), Nagorny (82850 people), located on the periphery of Kharkov, most of the hospitals providing emergency care are located in the center. These results can be used to optimize and develop the existing medical infrastructure and can be the subject of a separate study.

The obtained estimates of the accessibility level also do not contradict foreign studies [9], [10]: the greater the coverage of hospitals, the higher their level of accessibility. However, accounting the existing transport infrastructure and traffic congestion showed:

- hospitals are most accessible at 8 o'clock in the morning, least accessible – at 12 o'clock in the afternoon;
- in the same conditions of traffic congestion, an increase in the permissible delivery time increases the coverage area by an average of 3 times;

– the maximum coverage area is 219.9 km² (at 8:00 with a patient delivery time norm of 10 minutes).

At the same time, the widespread use of cartographic data and models made it possible to identify seven districts that are not included in the considered zone of accessibility of emergency hospitals. According to estimates, 98764 residents will not be able to receive medical care on time, which can also become a separate subject of research in the context of ongoing medical reform in Ukraine.

Correction of the network of existing medical centers on a model, a justified change in the position of only 3 hospitals led to the fact that the coverage area and the level of accessibility of hospitals increased significantly. For the 10-minute delivery zone, the access level is almost 97%, which significantly increases the number of people who will get timely medical care.

CONCLUSIONS

The current situation in the field of health has once again emphasized the need to improve and develop the medical sphere. Urban development projects should be aimed at finding effective ways to use the existing medical infrastructure and identifying areas for its improvement to further uplift the living conditions of the population. Therefore, of particular importance and relevance is the solution to the problem of the placement of medical facilities.

The scientific novelty: the method of choosing the location of the medical center as a geographic object has been improved, considering the number of patients, who need the services of the center, transport infrastructure and road congestion, which, unlike existing ones, allows making decisions based not only on the shortest distance from the center to the patient, but also on the level of its availability.

For the first time, a structure of information technology for multivariate analysis of a network of city medical centers is proposed, the use of which in decision support systems allows to form objective conclusions based on the results of the implementation of the method of spatial clustering using GIS technologies.

The practical significance of obtained results has been proven on a practical example – an analysis of the existing hospital system for the provision of emergency medical care in Kharkiv. Based on the assumptions made, it is shown that the level of access to medical care does not exceed 62%. As an option to improve the existing situation, a hypothetical example of improving the existing medical infrastructure is considered by correcting the network of existing medical centers and reasonably changing the position of some of them, which increases the level of accessibility by 1.5 times.

Prospects for further research – It is advisable to conduct further research in the direction of managing medical institution networks nationwide in order to optimize and rationalize the functioning of the medical industry in the country. This will increase the effectiveness of emergency departments, as well as help optimize the tasks of transplantology, where time is also a key factor.

This approach allows the most efficient distribution of budgetary funds for the creation of modern networks of medical infrastructure and monitoring the effectiveness of existing networks of medical institutions in order to improve them. It is obvious that geospatial analysis of the city's network infrastructure can also be applied not only in the field of healthcare, but also in any other areas related to interaction with a large number of people.

ACKNOWLEDGEMENTS

The results described in this article were obtained as part of the research work “Remote sensing data processing methodology for solving environmental monitoring tasks” “Methodology for processing Earth's remote sensing data for environmental monitoring tasks” (No. 0120U100530).

REFERENCES

1. Wilkerson M., Mitchell M., Shanahan D. et al. The role of socio-economic factors in planning and managing urban services, *Ecosystems Service*, 2018, Vol. 31, Part A, pp. 102–110.
2. Lopatkina A. E. Razmeshhenie objektov social'noj infrastruktury: zarubezhnyj opyt i vozmozhnosti dlja Rossijskoj praktiki [Electronic resource], *Russian Journal of Education and Psychology*, 2014, № 4 (36). Access mode: <https://www.sisp.nkas.ru>. DOI: 10.12731/2218-7405-2014-4-8.
3. Pomeranz J., Silver D. State Legislative Strategies to Pass, Enhance, and Obscure Preemption of Local Public Health Policy-Making [Electronic resource], *American Journal of Preventive Medicine*, 2020. Access mode: <https://www.sciencedirect.com/science/article/abs/pii/S074937972030180X>.
4. Social'no-politichna situacija v Ukraïni na pochatku grudnja 2019 roku [Electronic resource]. Access mode: <https://sosis.kiev.ua/2019-12-2>.
5. Smith K. E., Bambra C., Hill S. E. Health Inequalities: Critical Perspectives. USA, Oxford University Press, 2015, 318 p.
6. Artamonov I. V. Dostupnost' medicinskih uslug v regione: prostranstvennyj aspekt, *Sovremennaja jekonomika: problemy i reshenija*, 2019, № 10 (118), pp. 139–154. DOI: 10.17308/meps.2019.10/2233.
7. Lewis S., Willis K., Collyer F. Navigating and making choices about healthcare: The role of place, *Health & Place*, 2018, Vol. 52, pp. 215–220.
8. Dae Ko Y., Duk Song B., Hwang H. Location, capacity and capability design of emergency medical centers with multiple emergency diseases, *Computers & Industrial Engineering*, 2016, Vol. 101, pp. 10–20.
9. Zhang Y., Atkins D. Medical facility network design: User-choice and system-optimal models, *European Journal of Operational Research*, 2019, Vol. 273, Issue 1, pp. 305–309.
10. Sener I. N. Pendyala R. M., Bhart Ch. R. Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior, *Journal of Transport Geography*, 2011, Vol. 19, Issue 2, pp. 294–303.
11. Zhang Y., Liang L., Liu E. et al. Patient choice analysis and demand prediction for a health care diagnostics company, *European Journal of Operational Research*, 2016, Vol. 251, Issue 1, pp. 198–205.
12. Kharchenko S., Duzhiy I., Nechausov A. et al. A Population-Based Study with a Spatial Event Approach of Hepatobiliary Cancers in Ukraine, *Spring Fall Conference (KASL), 22–24 June 2017 : proceeding*. Korea, Korea Society of Liver Science, Vol. 1, pp. 181–182.
13. Vanolya N. M., Jelokhani-Niaraki M., Toomanian A. Validation of spatial multicriteria decision analysis results using public participation GIS [Electronic resource], *Applied Geography*, 2019, Vol. 112. Access mode: <https://www.sciencedirect.com/science/article/abs/pii/S0143622818310695>.
14. Kotowska D., Skorka P., Walasz K. Delineating the number of animal territories using digital mapping and spatial hierarchical clustering in GIS technology [Electronic resource], *Ecological Indicators*, 2019, Vol. 107. Access mode : <https://www.sciencedirect.com/science/article/abs/pii/S1470160X19306636>.
15. ArcGIS Desktop [Electronic resource]. Access mode: <https://desktop.arcgis.com/ru/arcmap/10.3/main/get-started/arcgis-tutorials.htm>.
16. Simpson S., Monroe Ch. Implementing and Evaluating a Standard of Care for Clinical Evaluations, *The Journal of Emergency Medicine*, 2018, Vol. 55, Issue 4, pp. 522–529.
17. Nechausov A., Kharchenko S. Epidemiologic assessment with a cohort study of current patient access to pancreatic transplantation in France, *The 31th Meeting of Japanese Society of Hepato-Biliary-Pancreatic Surgery, 15–19 June 2019 : proceeding*. Kagawa, Kagawa university, 2019, P. 331.

Received 22.07.2020.
Accepted 19.08.2020.

ВИРІШЕННЯ ПРОБЛЕМИ РОЗМІЩЕННЯ МЕДИЧНИХ ОБ'ЄКТІВ У ПРОЕКТАХ РОЗВИТКУ МІСТ

Даншина С. Ю. – канд. техн. наук, доцент, доцент кафедри геоінформаційних технологій та космічного моніторингу Землі в Національному аерокосмічному університеті ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Нечаусов А. С. – канд. техн. наук, доцент кафедри геоінформаційних технологій і космічного моніторингу Землі в Національному аерокосмічному університеті ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

АНОТАЦІЯ

Актуальність. Розглянуто проблему раціонального розміщення медичних установ населеного пункту. Запропоновано методи оцінки ефективності використання існуючої медичної мережі та пошуку шляхів її покращення при реалізації проєктів міського розвитку. Метою роботи є побудова і вивчення порядку вирішення проблеми розміщення медичних установ з урахуванням існуючої інфраструктури міста для виконання вимог доступності.

Мета. Метою роботи є побудова та вивчення процедури вирішення проблеми розміщення медичних закладів із врахуванням існуючої інфраструктури міста для виконання вимог щодо доступності.

Метод. Сформовано набір факторів, які впливають на розміщення медичних установ і дозволяють прийняти систематичне і раціональне рішення при виборі місця. Запропоновано метод вирішення проблеми розміщення медичних установ, що забезпечує підвищення рівня їх доступності з урахуванням чисельності населення, його просторового розподілу, існуючої в населеному пункті транспортної розв'язки і пробок на дорогах. Процедура вирішення завдання представлена у вигляді моделі IDEF0. Метод заснований на геоінформаційному аналізі даних, результати якого шляхом просторової кластеризації представлені у вигляді набору картографічних моделей, сукупність яких дозволяє сформулювати рішення про місце розташування. Метод дозволяє оцінити рівень доступності існуючих медичних установ в населеному пункті, створити список місць для їх можливого розташування, виявити райони місцевості, що не входять в зону доступності.

Отримані результати. Удосконалено метод вибору місця розташування медичного центру, що дає можливість прийняти рішення, спираючись не лише на найкоротші відстані від центру до пацієнта, а й на рівні його доступності. Вперше запропоновано структуру інформаційної технології багатфакторного аналізу мережі медичних центрів міста для систем підтримки прийняття рішень за допомогою ГІС-технологій.

Висновки. Проблема розміщення медичних установ, яка полягає в підвищенні рівня доступності за рахунок розширення зони медичного охоплення території, на якій проживає населення міста, з використанням ГІС, була вирішена практично і представлена у вигляді моделі IDEF0, яка визначає процедуру вирішення проблеми і вибору місця розташування медичного центру. Практична значимість отриманих результатів була доведена на практичному прикладі – аналізі існуючої системи закладів надання невідкладної медичної допомоги в Харкові.

КЛЮЧОВІ СЛОВА: рівень доступності, зона охоплення, населення, час доставки пацієнта, модель IDEF0, геоінформаційні технології, мережевий аналіз, ArcGIS.

РЕШЕНИЕ ПРОБЛЕМЫ РАЗМЕЩЕНИЯ МЕДИЦИНСКИХ ОБЪЕКТОВ В ПРОЕКТАХ РАЗВИТИЯ ГОРОДОВ

Даншина С. Ю. – канд. техн. наук, доцент, доцент кафедры геоинформационных технологий и космического мониторинга Земли в Национальном аэрокосмическом университете им. Н. Е. Жуковского «Харьковский авиационный институт», Харьков, Украина.

Нечаусов А. С. – канд. техн. наук, доцент кафедры геоинформационных технологий и космического мониторинга Земли в Национальном аэрокосмическом университете им. Н. Е. Жуковского «Харьковский авиационный институт», Харьков, Украина.

АННОТАЦИЯ

Актуальность. Рассмотрена проблема рационального размещения медицинских учреждений населенного пункта. Предложены методы оценки эффективности использования существующей медицинской сети и поиска путей ее улучшения при реализации проектов городского развития. Целью работы является построение и изучение порядка решения проблемы размещения медицинских учреждений с учетом существующей инфраструктуры города для выполнения требований доступности.

Цель. Целью работы является построение и изучение порядка решения проблемы размещения медицинских учреждений с учетом существующей инфраструктуры города для выполнения требований доступности.

Метод. Сформирован набор факторов, которые влияют на размещение медицинских учреждений и позволяют принять систематическое и рациональное решение при выборе места. Предложен метод решения проблемы размещения медицинских учреждений, обеспечивающий повышение уровня их доступности с учетом численности населения, его пространственного распределения, существующей в населенном пункте транспортной развязки и пробок на дорогах. Процедура решения задачи представлена в виде модели IDEF0. Метод основан на геоинформационном анализе данных, результаты которого путем пространственной кластеризации представлены в виде набора картографических моделей, совокупность которых позволяет сформировать решение о местоположении. Метод позволяет оценить уровень доступности существующих медицинских учреждений в населенном пункте, создать список мест для их возможного расположения, обнаружить районы местности, не входящие в зону доступности.

Полученные результаты. Усовершенствован метод выбора места расположения медицинского центра, дающий возможность принятия решений, основываясь не только на кратчайшем расстоянии от центра до пациента, но и на уровне его доступности. Впервые предложена структура информационной технологии многофакторного анализа сети медицинских центров города для систем поддержки принятия решений с помощью ГИС-технологий.

Выводы. Проблема размещения медицинских учреждений, которая заключается в повышении уровня доступности за счет расширения зоны медицинского охвата территории, на которой проживает население города, с использованием ГИС, была решена практически и представлена в виде модели IDEF0, которая определяет процедуру решения проблемы и выбора расположения медицинского центра. Практическая значимость полученных результатов была доказана на практическом примере – анализе существующей системы учреждений оказания неотложной медицинской помощи в Харькове.

КЛЮЧЕВЫЕ СЛОВА: уровень доступности, зона охвата, население, время доставки пациента, модель IDEF0, геоинформационные технологии, сетевой анализ, ArcGIS.

ЛИТЕРАТУРА / LITERATURA

1. The role of socio-economic factors in planning and managing urban services / [M. Wilkerson, M. Mitchell, D. Shanahan et al.] // *Ecosystems Service*. – 2018. – Vol. 31, Part A. – P. 102–110.
2. Лопаткина А. Е. Размещение объектов социальной инфраструктуры: зарубежный опыт и возможности для Российской практики [Электронный ресурс] / А. Е. Лопаткина // *Russian Journal of Education and Psychology*. – 2014. – № 4 (36). – Режим доступа: <https://www.sisp.nkas.ru>. DOI: 10.12731/2218-7405-2014-4-8.
3. Pomeranz J. State Legislative Strategies to Pass, Enhance, and Obscure Preemption of Local Public Health Policy-Making [Electronic resource] / J. Pomeranz, D. Silver // *American Journal of Preventive Medicine*. – 2020. – Access mode: <https://www.sciencedirect.com/science/article/abs/pii/S074937972030180X>.
4. Соціально-політична ситуація в Україні на початку грудня 2019 року [Електронний ресурс]. – Режим доступу: <https://socis.kiev.ua/2019-12-2>.
5. Smith K. E. Health Inequalities: Critical Perspectives / K. E. Smith, C. Vambra, S. E. Hill. – USA : Oxford University Press, 2015. – 318 p.
6. Артамонов И. В. Доступность медицинских услуг в регионе: пространственный аспект / И. В. Артамонов // *Современная экономика: проблемы и решения*. – 2019. – № 10 (118). – С. 139–154. DOI: 10.17308/meps.2019.10/2233.
7. Lewis S. Navigating and making choices about healthcare: The role of place / S. Lewis, K. Willis, F. Collyer // *Health & Place*. – 2018. – Vol. 52. – P. 215–220.
8. Dae Ko Y. Location, capacity and capability design of emergency medical centers with multiple emergency diseases / Y. Dae Ko, B. Duk Song, H. Hwang // *Computers & Industrial Engineering*. – 2016. – Vol. 101. – P. 10–20.
9. Zhang Y. Medical facility network design: User-choice and system-optimal models / Y. Zhang, D. Atkins // *European Journal of Operational Research*. – 2019. – Vol. 273, Issue 1. – P. 305–309.
10. Sener I. N. Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior / I. N. Sener, R. M. Pendyala, Ch. R. Bhart // *Journal of Transport Geography* – 2011. – Vol. 19, Issue 2. – P. 294–303.
11. Patient choice analysis and demand prediction for a health care diagnostics company / [Y. Zhang, L. Liang, E. Liu et al.] // *European Journal of Operational Research*. – 2016. – Vol. 251, Issue 1. – P. 198–205.
12. A Population-Based Study with a Spatial Event Approach of Hepatobiliary Cancers in Ukraine / [S. Kharchenko, I. Duzhiy, A. Nechausov et al.] // *Spring Fall Conference (KASL), 22–24 June 2017 : proceeding*. – Korea: Korea Society of Liver Science. – Vol. 1. – P. 181–182.
13. Vanolya N. M. Validation of spatial multicriteria decision analysis results using public participation GIS [Electronic resource] / N. M. Vanolya, M. Jelokhani-Niaraki, A. Toomanian // *Applied Geography*. – 2019. – Vol. 112. – Access mode: <https://www.sciencedirect.com/science/article/abs/pii/S0143622818310695>.
14. Kotowska D. Delineating the number of animal territories using digital mapping and spatial hierarchical clustering in GIS technology [Electronic resource] / D. Kotowska, P. Skorka, K. Walasz, // *Ecological Indicators*. – 2019. – Vol. 107. – Access mode : <https://www.sciencedirect.com/science/article/abs/pii/S1470160X19306636>.
15. ArcGIS Desktop [Electronic resource]. – Access mode: <https://desktop.arcgis.com/ru/arcmap/10.3/main/get-started/arcgis-tutorials.htm>.
16. Simpson S. Implementing and Evaluating a Standard of Care for Clinical Evaluations / S. Simpson, Ch. Monroe // *The Journal of Emergency Medicine*. – 2018. – Vol. 55, Issue 4. – P. 522–529.
17. Nechausov A. Epidemiologic assessment with a cohort study of current patient access to pancreatic transplantation in France / A. Nechausov, S. Kharchenko // *The 31th Meeting of Japanese Society of Hepato-Biliary-Pancreatic Surgery, 15–19 June 2019 : proceeding*. – Kagawa : Kagawa university, 2019. – P. 331.

MODIFIED ALGORITHM FOR SEARCHING THE ROOTS OF THE ERROR LOCATORS POLYNOMIAL WHILE DECODING BCH CODES

Krylova V. A. – PhD, Associate Professor of the Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkov, Ukraine.

Tverytnykova E. E. – Dr. Sc., Professor of the Department of information and measuring technologies and systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine.

Vasylchenkov O. G. – PhD, Associate Professor of the Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkov, Ukraine.

Kolisnyk T. P. – PhD, Associate Professor of the Department of information technology and cybersecurity, Kharkiv National University of Internal Affairs, Kharkov, Ukraine.

ABSTRACT

Context. In telecommunications and information systems with an increased noise component the noise-resistant cyclic BCH and Reed-Solomon codes are used. The adjustment and correcting errors in a message require some effective decoding methods. One of the stages in the procedure of decoding RS and BCH codes to determine the position of distortions is the search for the roots of the error locator polynomial. The calculation of polynomial roots, especially for codes with significant correction capacity is a laborious task requiring high computational complexity. That is why the improvement of BCH and RS codes decoding methods providing to reduce the computational complexity is an urgent task.

Objective. The investigation and synthesis of the accelerated roots search algorithm of the error locator polynomial presented as an affine polynomial with coefficients in the finite fields, which allows accelerating the process of BCH and RS code decoding.

Method. The classical roots search method based on the Chan's algorithm is performed using the arithmetic of the Galois finite fields and the laborious calculation, in this case depends on the number of addition and multiplication operations. For linearized polynomials, the roots search procedure based on binary arithmetic is performed taking into account the values obtained at the previous stages of the calculation, which provides the minimum number of arithmetic operations.

Results. An accelerated algorithm for calculating the values of the error locator polynomial at all points of the $GF(2^m)$ finite field for linearized polynomials based on the Berlekamp-Massey method has been developed. The algorithm contains a minimum number of addition operations, due to the use at each stage of the calculations the values obtained at the previous step, as well as the addition in the finite field $GF(2)$. A modified roots search method for affine polynomials over the finite fields has been proposed to determine error positions in the code word while decoding the cyclic BCH and RS codes.

Conclusions. The scientific newness of the work is to improve the algorithm of calculating the roots of the error locator polynomial, which coefficients belong to the elements of the finite field. At the same time it simplifies the procedure for cyclic BCH and RS codes decoding, due to reducing the computational complexity of one of the decoding stages, especially finding the error positions using the modified Berlekamp-Massey algorithm. These facts are confirmed by the simulation program results of the roots search of the error locator polynomial algorithm. It is shown, that the application of the accelerated method permits to reach a gain on speed of 1.5 times.

KEYWORDS: BCH codes, error locator polynomial, Chan's search, Berlekamp-Massey algorithm and Reed-Solomon codes.

ABBREVIATIONS

BM is a Berlekamp-Massey;
BCH is a Bose-Chaudhuri-Hocquenghem;
RS is a Reed-Solomon.

σ_0 is a zero coefficient of locators of errors;
 σ_i is an element of the finite field $GF(2^m)$.

NOMENCLATURE

A is a binary matrix of linearized polynomials;
 $F(x)$ is a linearized polynomial;
 f_i is a coefficient to the $GF(2^m)$ finite field;
 $GF(2^m)$ is a finite field Galois;
 l is a field order;
 m is a natural number;
 $p(x)$ is a generating polynomial;
 t is a times;
 Y is a binary vector of zero coefficient of the error locator polynomial;
 v is a degree of the errors locators polynomial;
 α^j is an element of $GF(2^m)$ finite field;
 β^j is a binary vector field element α ;
 $\sigma(x)$ is an errors locators polynomial;

INTRODUCTION

One of the ways to protect information from errors in digital communication systems is to use error-correcting codes detecting and correcting errors in the information transmission channel. The requirements for encoding and decoding methods and procedures by the reference to the spectral and energy efficiency of a communication system give the task of constructing simplified algorithms for correcting errors in transmitted information. In modern information systems, cyclic BCH and Reed-Solomon codes, which require high redundancy, are the most used to ensure their corrective abilities. Moreover, the processing time of information in a decoding device, which depends on the complexity of the encoding and decoding algorithms, limits significantly the operating time of the error protection system.

BCH and Reed-Solomon codes used in modern information storage and transmission systems built on classical error detection and correction procedures, at some decoding stages, due to a significant amount of data, have high arithmetic complexity. This leads to limited potential possibilities of the above said correction codes to provide a given probability of information loss.

In this case, in order to increase the operation speed of information protection systems from errors in information data transmission systems, it is necessary to improve and refine existing algorithms for decoding of cyclic BCH codes for improving the correcting ability of the code by reducing the computational complexity of encoding and decoding procedures with high redundancy.

The object of research is the decoding process of cyclic RS and BCH codes in the digital communication systems.

The subject of research is a method for finding the roots of the error locator polynomial whose coefficients belong to in the Galois fields.

The aim of this work is to study the decoding methods of RS and BCH codes, allowing detecting and correcting errors in the code sequence, as well as the development of an accelerated algorithm for calculating the roots of the error polynomial, the implementation of the algorithm into Visual Studio and the subsequent analysis of the results.

1 PROBLEM STATEMENT

The ways of BCH and RS codes decoding are quite well developed in theory and practice, which is presented in [1], but nevertheless the implementation of decoding algorithms is quite a laborious task, especially if the finite fields of a large order are used.

A typical procedure for the decoding of cyclic RS codes is proposed by the author R.E. Blahut [2] and consists of the following stages:

- the calculation of syndrome components (syndrome vector);
- the formation of a key equation and finding of the error locator polynomial by one of the methods – Peterson, Berlekamp-Massey or the Euclidean algorithm;
- the searching for the roots of the error locator polynomial using the Chan's method – a complete enumeration of all values;
- the calculation of the error values polynomial and error character determination based on the Forney algorithm;
- the correction of erroneous characters.

The $GF(2^m)$ finite field contains 2^m elements $(1, \alpha^1, \alpha^2, \alpha^3, \dots)$, each of which is represented by a binary vector of m bits and in the practice PC code mostly uses the calculations in the $GF(2^m)$ finite fields.

One of the most time-consuming stages of RS code decoding is searching for the roots of the error locator polynomial

$$\sigma(x) = \sigma_0 + \sigma_1 x + \sigma_2 x^2 + \dots + \sigma_v x^v. \quad (1)$$

It is known that the Chan's method can be used to find the roots of the error locator polynomial (1) degree v in the field $GF(2^m)$. However, the Chan's algorithm requires multiplication of each coefficient σ_i in the formula (1) by the element of field $GF(2^m)$ using α degrees.

Therefore, to correct errors in the code sequence, the Chan's method is used, which in fact is a complete enumeration of all elements, which significantly affects the effectiveness of the decoding device. If the decoder has a

known error locator polynomial $\sigma(x) = \sum_{i=0}^t \sigma_i x^i$, which

roots are mutual to the error position locators, then the Chan's procedure can be applied to each of the locators $1, \alpha^1, \alpha^2, \alpha^3, \dots, \alpha^l$ ($l = 2^m - 1$) to check if the symbol displayed at the moment is an error [1]. I.e., for all non-zero elements of α^k field $GF(2^m)$, the condition $\sigma(\alpha^k) = 0$ is checked, and its implementation indicates that α^k is a mutual root of the error locator polynomial to the erroneous symbol. In this case, the number of operations to implement the Chan's procedure is $2t(2^m - 1)$.

Thus, the Chan's procedure for searching the roots of the polynomial error becomes quite complicated for calculations in large finite fields ($m > 8$) and for error locators polynomials of a large degree, because it requires a significant number of operations.

Therefore, the task of this work is to improve the method for determining the positions of distortions in the code word and to reduce the computational complexity of the algorithm for finding the roots of the error polynomial in the finite fields of $GF(2^m)$ when decoding RS and BCH codes.

2 REVIEW OF THE LITERATURE

The principles of encoding, more modern methods of errors control, code applications for the design of real error control systems are covered in [3]. With the help of a special class of p -polynomials the authors offered a fast algorithm for finding polynomial roots of the degree less than 5. In this algorithm the polynomial defined in the formula (1) is transformed into an affine polynomial above the $GF(2^m)$ field. Then the roots can be found by solving a system from m linear equations with m unknown above the $GF(2)$ field. Therefore, at the minor values of v , the advantage of the above algorithm over the Chan's method is that the calculations in the $GF(2^m)$ field (multiplying σ_i coefficients by α degrees) required in the standard Chan's search method are completely excluded.

The searching method for the roots of the polynomial on the base of some transformation that makes it possible to group some polynomial constituents of the degree no higher than 11 into multiple affine polynomials is proposed in [4]. This method makes it possible to accelerate the calculations, but this algorithm has its disadvantages. For example, it can be used for polynomials which degrees cannot be higher than 11. The building and decoding of BCH codes for different communication systems are described in [5]. The authors propose an improved algorithm for searching the roots of polynomials over the

finite fields. This algorithm significantly accelerates the process of BCH codes decoding. Also the improved hybrid algorithm of polynomial roots finding over finite fields is considered in [6] and [7] articles. The authors have combined the Jiang's algorithm, which is based on the decomposition of the error locator polynomial in the sum of multiple affine polynomials with modified analytical methods for solving polynomials with small degree in radicals. A number of works have been devoted to different issues of BCH and RS decoding, where simple equation for Reed-Solomon codes on the base of uniting of two algorithms (Berlekamp-Massey algorithm and Euclidean) was proposed [8], the new approach to calculating the total error locator polynomial [9–10], the dependence of indicators of reliability and efficiency of information transmission on the parameters of Reed-Solomon codes are given in [11], the main stages of Reed-Solomon codes encoding (decoding) for practical application [12] were reviewed.

The analysis of publications has shown the necessity to study the procedures of cyclic BCH codes decoding, as well as the methods of determining error positions in the code word, based on the search of the roots of the error locator polynomial. The calculations of the roots values are carried out using the elements of finite fields, taking into account the degree of error polynomial, which significantly affects the calculation time. Therefore, the analysis results of known methods of roots search will be presented further, and also the accelerated algorithms of definition of errors positions in a code word at decoding RS and BCH codes will be described.

3 MATERIALS AND METHODS

The solution of the problem of searching the roots of the error locator polynomial (1), which σ_i coefficients belong to the $GF(2^m)$ finite field, using the Berlekamp algorithm is based on a special class of polynomials. These polynomials, which roots can be found much easier, are called p -polynomials or linearized polynomials [1].

$$F(x) = \sum_i f_i x^{2^i} . \quad (2)$$

If we suggest that the error locators polynomial has the degree $v=2^l$, then it can be represented as

$$\sigma(x) = F(x) + \sigma_0, \quad \sigma_0 \in GF(2^m) . \quad (3)$$

The polynomial of the type (3) is called affine polynomial [6].

If the polynomial is checked when searching for roots, then we obtain

$$F(x) = \sigma_0, \quad \sigma_0 \in GF(2^m) . \quad (4)$$

For any $GF(2^m)$ finite field the standard basis is a set of m elements $1, \alpha^1, \alpha^2, \alpha^3, \dots, \alpha^{m-1}$ [14]. The values of the linearized polynomial $F_0(\alpha^0) F_1(\alpha^1) \dots F_{m-1}(\alpha^{m-1})$ at the points of the standard basis of the $GF(2^m)$ finite field can be calculated using (2). The obtained result lies in the $GF(2^m)$ field, but for further calculations it is reasonable to present it as a binary vector corresponding to the field element

$$F_i(\alpha^i) = \{a_0 a_1 a_2 \dots a_{m-1}\}, \quad a_k \in GF(2) .$$

We denote the searching root of the error polynomial as $\alpha \in GF(2^m)$ field element through a binary vector

$$\beta^i = \{b_0 b_1 b_2 \dots b_{m-1}\}, \quad b_k \in GF(2) .$$

Then it is derived from the formula (1) that the zero coefficient of error locator polynomials represented as $Y = \{y_0 y_1 \dots y_{m-1}\}$ binary vector can be obtained by multiplying $\beta^i = \{b_0 b_1 b_2 \dots b_{m-1}\}, b_k \in GF(2)$ the vector of the line by a matrix A over field $GF(2^m)$

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{m-1} \end{bmatrix} \begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} & \dots & a_{0,m-1} \\ a_{1,0} & a_{1,1} & a_{1,2} & \dots & a_{1,m-1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m-1,0} & a_{m-1,1} & \dots & a_{m-1,m-1} \end{bmatrix} = [y_0 y_1 \dots y_{m-1}] . \quad (5)$$

Therefore, the decomposition coefficients of $F(x)$ polynomial on a standard basis can be obtained

$$F(x) = \sum_{k=0}^{m-1} b_k F(\alpha^k), \quad b_k \in GF(2) . \quad (6)$$

For example, it is necessary to find the roots of the error locator polynomial that coefficients belong to the $GF(2^3)$ finite field, generated by the primitive polynomial $p(x)=x^3+x+1$:

$$\sigma(x) = x^2 + \alpha^6 x + \alpha^6 . \quad (7)$$

We represent the error polynomial as $\sigma(x) = F(x) + \sigma_0$, where the $F(x) = x^2 + \alpha^6 x$ linearized polynomial above the $GF(2^3)$ field and the zero coefficient is $\sigma_0 = \alpha^6$. Then, in accordance with (4), we obtain

$$F(x) = x^2 + \alpha^6 x = \alpha^6 . \quad (8)$$

The element of the $GF(2^3)$ α^6 finite field is comparable to the $[y_2 y_1 y_0]$ binary vector. If the $\beta^i = \{b_0 b_1 b_2\}, b_k \in GF(2)$ searching roots of the error locator polynomial, then in accordance with (6) we obtain

$$b_0F(\alpha^0) + b_1F(\alpha^1) + b_2F(\alpha^2) = [y_0 \ y_1 \ y_2]. \quad (9)$$

The basis vectors $F_2(\alpha^2)$, $F_1(\alpha^1)$ и $F_0(\alpha^0)$ taking into account (8) can be calculated as::

$$\begin{aligned} F_0(\alpha^0) &= \alpha^0 + \alpha^6 = \alpha^2, \\ F_1(\alpha^1) &= \alpha^2 + \alpha^7 = \alpha^6, \\ F_2(\alpha^2) &= \alpha^4 + \alpha^1 = \alpha^2. \end{aligned}$$

Each $F(\alpha^i)$ basis vector above the $GF(2^3)$ field corresponds to a binary

$$\begin{aligned} F(\alpha^i) &= \{a_0 \ a_1 \ a_2\}, \quad a_k \in GF(2) \\ F(\alpha^0) &= \{1 \ 0 \ 0\}, \quad F(\alpha^1) = \{1 \ 0 \ 1\}, \quad F(\alpha^2) = \{1 \ 0 \ 0\}. \end{aligned}$$

Then, in accordance with (5), we have a system of the linear binary equations

$$[b_0 \ b_1 \ b_2] \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = [1 \ 0 \ 1].$$

Therefore, the roots of the error locator polynomial (7) can be found by solving simultaneously three equations with three unknown $\{b_0 \ b_1 \ b_2\}$, $b_k \in GF(2)$. It is obvious that the elements of the $GF(2^3)$ finite field $\alpha^1 \{0 \ 1 \ 0\}$ and $\alpha^5 \{1 \ 1 \ 1\}$ are the roots of the polynomial (7). This example shows that to find the roots of the polynomial (7) by means of the Berlekamp algorithm it is necessary to calculate only three $F(\alpha^2)$, $F(\alpha^1)$ and $F(\alpha^0)$ in the $GF(2^3)$ field instead of calculating all values $F(\alpha^0)$, $F(\alpha^1)$... $F(\alpha^6)$, required for searching the roots by the Chan's method.

For the decrease in complexity and quantity of computational operations at finding the roots of errors locator polynomial by the Berlekamp method it is expedient to use earlier calculated values at the following stage of algorithm. For this purpose we will use the property of the linearized polynomial

$$F(x + y) = F(x) + F(y). \quad (10)$$

Thus, in accordance with the property (10), we obtain

$$F(\alpha^i + \alpha^j) = F(\alpha^i) + F(\alpha^j) \quad (11)$$

or

$$F(\alpha^k) + F(\alpha^i) = F(\alpha^j), \quad \alpha^k = \alpha^i + \alpha^j.$$

Consider such a pair of α^k and α^i elements of the finite field $GF(2^m)$ as a standard basis (binary vector of m length), which would differ from each other in one position. Then in the formula (11) the α^j field element is in this case a single vector of m length. As it has been pointed

ed out earlier such elements (binary sequence weight 1) are basic in the finite field $GF(2^m) \square 1, \alpha^1, \alpha^2, \alpha^3, \dots, \alpha^{m-1}$.

Hence, if we rank (order) the elements of the $GF(2^m)$ finite field in such a way that the nearby vectors will differ exactly in one position. Then, at each step of the algorithm for searching the roots of the error locator polynomial, the calculations are reduced to a single addition \square the previous value and the value of the linearized polynomial at the points of the standard basis $F(\alpha^0) F(\alpha^1) \dots F(\alpha^{m-1})$ of the $GF(2^m)$ finite field.

Thus, in order to calculate all values of the error locator polynomial represented as the affine polynomial (3) in all points of the finite field $\beta^i \in GF(2^m)$, $i = 0, 1, \dots, m-1$, it is necessary to complete the following steps

$$F(\beta^i) = F(\beta^{i-1}) + F_j(\alpha^j), \quad (12)$$

($\alpha^j = \beta^i \oplus \beta^{i-1}$ – corresponds to one of the basic elements of the $GF(2^m)$ field).

The expression (12) provides the procedure for finding the polynomial value set $\sigma(x)$ at all points $\alpha^i \in GF(2^m)$. The calculation requires ordering of all the elements of the field, the presence of the previous value $F(\beta^{i-1})$ and the previously found values of the basis vectors $F(\alpha^i)$.

Thus, to find the roots of the error locator polynomial, and therefore to calculate the polynomial value (3) at all points of the finite field it is necessary to use the following algorithm:

1. To set the value of zero coefficient of error locators polynomial σ_0 and linearized polynomial $F(x)$.
2. To find the values of the linearized polynomial $F(x)$ at the points of the standard basis of the finite field $GF(2^m)$

$$F_k = F(\alpha^i), \quad i = 0, 1, \dots, m-1.$$

3. To perform the initialization $F(\beta^0) = 0$.
4. To arrange all elements of the $\alpha^i \in GF(2^m)$ finite field in the form of binary vectors in such a way that two any nearby vectors will differ exactly in one position.
5. To calculate the value of

$$F(\beta^i) = F(\beta^{i-1}) + F_j(\alpha^j), \quad \alpha^j = \beta^i \oplus \beta^{i-1}.$$

6. If $F(\beta^i) = \sigma_0$, then the field element β^i is the polynomial root of $\sigma(x)$.

4 EXPERIMENTS

To confirm the functionality of the above-mentioned theoretical calculations we will perform the procedure of searching the roots of the error locator polynomial, which coefficients belong to the $GF(2^4)$ finite field, generated by the primitive polynomial $p(x) = x^4 + x + 1$:

$$\sigma(x) = x^2 + \alpha^5 x + 1. \quad (13)$$

1. We represent the error polynomial (13) as $\sigma(x) = F(x) + \sigma_0$, where the linearized polynomial $F(x) = x^2 + \alpha^5 x$ above the $GF(2^4)$ and the zero coefficient $\sigma_0 = 1$.

2. We will find the values of the linearized polynomial at the points of the standard basis of the $GF(2^4)$ finite field

$$F_0(\alpha^0) = \alpha^{10} (1010), F_1(\alpha^1) = \alpha^5 (1011), F_2(\alpha^2) = \alpha^8 (1110), F_3(\alpha^3) = \alpha^0 (0001).$$

3. We will perform the initialization $F(\beta^0) = 0$.

4. The elements of the $GF(2^4)$ finite field, decomposed as the binary vectors, should be ordered so that the nearby vectors will differ in one position.

Table 1 – The $GF(2^4)$ Finite Field Element Ordering

Field Element	Binary Vector	Field Element	Binary Vector
$\beta^0(\alpha^0)$	0000	$\beta^8(\alpha^{13})$	1100
$\beta^1(\alpha^1)$	0001	$\beta^9(\alpha^{13})$	1101
$\beta^2(\alpha^{12})$	0011	$\beta^{10}(\alpha^{13})$	1111
$\beta^3(\alpha^1)$	0010	$\beta^{11}(\alpha^{13})$	1110
$\beta^4(\alpha^{13})$	0110	$\beta^{12}(\alpha^{13})$	1010
$\beta^5(\alpha^7)$	0111	$\beta^{13}(\alpha^{13})$	1011
$\beta^6(\alpha^{13})$	0101	$\beta^{14}(\alpha^{13})$	1001
$\beta^7(\alpha^{13})$	0100	$\beta^{15}(\alpha^{13})$	1000

5. We will perform step-by-step calculations in accordance with (12).

1) $\beta^0 \oplus \beta^1 = 0001$ – corresponds to the basic element α^0 the $GF(2^4)$ field

$$F(\beta^1) = F(\beta^0) + F_0(\alpha^0) = 0000 + 1010 = 1010;$$

2) $\beta^1 \oplus \beta^2 = 0010$ – corresponds to the basic element α^1 the $GF(2^4)$ field

$$F(\beta^2) = F(\beta^1) + F_1(\alpha^1) = 1010 + 1011 = 0001;$$

3) $\beta^2 \oplus \beta^3 = 0001$ – corresponds to the basic element α^0 the $GF(2^4)$ field

$$F(\beta^3) = F(\beta^2) + F_0(\alpha^0) = 0001 + 1010 = 1011;$$

4) $\beta^3 \oplus \beta^4 = 0100$ – corresponds to the basic element α^2 the $GF(2^4)$ field

$$F(\beta^4) = F(\beta^3) + F_2(\alpha^2) = 1011 + 1110 = 0101;$$

...

14) $\beta^{13} \oplus \beta^{14} = 0010$ – corresponds to the basic element α^1 the $GF(2^4)$ field

ment α^1 the $GF(2^4)$ field

$$F(\beta^{14}) = F(\beta^{13}) + F_1(\alpha^1) = 0000 + 1011 = 1011;$$

15) $\beta^{14} \oplus \beta^{15} = 0001$ – corresponds to the basic element α^0 the $GF(2^4)$ field

$$F(\beta^{15}) = F(\beta^{14}) + F_0(\alpha^0) = 1011 + 1010 = 0001.$$

As $F(\beta^2) = 1$ и $F(\beta^{15}) = 1$, the roots of the polynomial (13) are consequently the elements of the field β^2 and β^{15} , α^{12} and α^3 respectively.

To estimate the real efficiency of the proposed modified algorithm for calculating the roots of the error locator polynomial, the software simulation in C++ language has been implemented. The multiplication in the $GF(2^8)$ finite fields for the Chan's method was carried out using the tables of logarithms and antilogarithms. The calculations of the roots and the comparison of results were carried out only for the linearized error polynomials and for the elements of the $\alpha^0, \alpha^1, \dots, \alpha^{254}$ field.

5 RESULTS

Figure 1 represents the measurement results of the computation speed for the modified root search method and for the Chan's method.

Based on the results, it can be concluded that the application of the modified algorithm for searching the roots of error locator polynomials, presented as the linearized polynomials, makes it possible to achieve a speed gain of 1.5 times in comparison with the Chan's method.

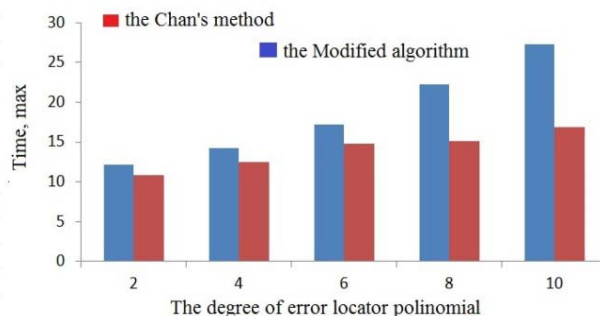


Figure 1 – The speed of the error polynomial roots computation

For the $GF(2^8)$ finite field, the average number of operations to search for the of polynomial roots of 8 degree is 4100 for the Chan's procedure and 1532 for the proposed modified root search method. For 16 degree of polynomial, the average number of operations is 8240 and 3562 respectively.

6 DISCUSSION

As all polynomials of the second degree are linearized, their roots can be found by solving the corresponding system of linear equations by the Berlekamp-Massey method [7]. Also, a linearized polynomial of error loca-

tors with a nonzero coefficient can be represented as an affine polynomial. In this case, the root search procedure is reduced to the addition of binary vectors in the GF(2) field GF as the elements of the GF(2^m) finite field, which allows reducing significantly the number of addition and multiplication operations of the field elements.

If each element α^i of the GF(2^m) finite field can be compared to a binary vector of m length, the field elements can always be ordered in such a way that two nearby vectors will differ from each other exactly in the same position, for example, using the Gray code.

The described algorithm allows finding the roots of the error locator polynomial if the polynomial (1) is represented as an affine polynomial (3). However, this root-searching procedure can also be generalized in the case of arbitrary polynomials if they are broken down into a set of affine polynomials.

CONCLUSIONS

In this work the acute scientific task of the accelerated search of the roots of the error locator polynomial which define the errors positions in the accepted code sequence at the stage of cyclic BCH and RS codes decoding is solved.

Also the effective algorithm of the roots calculation of the errors polynomial above the GF(2^m) finite fields is developed. The improved algorithm for calculating the roots of the polynomial errors with the coefficients in the finite field based on the Berlekamp-Massey algorithm for the linearized polynomials has been proposed, which provides a minimum number of arithmetic operations due to the use of data obtained from the previous stages of calculations. The proposed algorithm reduces the complexity of root calculations at one point of the finite field due to the application of a special arrangement of all elements of the finite field.

Practical newness of the work results lies in the implementation of the program model of the code message decoding system, which allows conducting the error correction, both for the classical method of searching the roots, and for the developed modified algorithm for calculating the positions of the disturbed characters. The modeling results have confirmed the efficiency of the proposed algorithm for the root searching of the error locator polynomial while the cyclic BCH and RS codes decoding, which allows achieving the speed gain by 1.5 times, compared with the Chan's search method.

ACKNOWLEDGEMENTS

The work is supported by National Technical University of Ukraine «Kharkiv Polytechnic Institute».

REFERENCES

1. Berlekamp E. R. Algebraic Coding Theory. New York, McGraw-Hill, 1968, 466 p.

2. Blahut Richard E. Algebraic Codes for Data Transmission. Cambridge, Cambridge University Press, 2003, 482 p.
3. Lin S., Costello D. J. Error control coding: fundamentals and applications. Prentice-Hall Inc, Printed in the United States of America, 2004, 624 p.
4. Truong T. K., Jeng J. H., Reed I. S. Fast algorithm for computing the roots of error locator polynomials up to degree 11 in Reed-Solomon decoders, *IEEE Transactions on Communications*, 2001, Vol. 49, Issue 5, pp. 779–783. DOI:10.1109/26.923801.
5. Nabipour S., Javidan J., Zare Gholamreza F. Error Detection Mechanism Based on Bch Decoder and Root Finding of Polynomial Over Finite Fields, *Journal of Mathematics and Computer Science*, 2014, Issue 4, pp. 271–281. DOI: <http://dx.doi.org/10.22436/jmcs.012.04.03>.
6. Fedorenko S. V., Trifonov P. V. Finding roots of polynomials over finite fields, *IEEE Transactions on Communications*, 2002, Vol. 50, Issue 11, pp. 1709–1711. DOI: 10.1109/TCOMM.2002.805269.
7. Fedorenko S., Trifonov P., Costa E., Haas H. Improved hybrid algorithm for finding roots of error-locator polynomials, *European Transactions on Telecommunications*, 2003, Vol. 14, Issue 5, pp. 411–416. DOI: <https://doi.org/10.1002/ett.936>.
8. Bras-Amorós M., Michael O'Sullivan E. The Symmetric Key Equation for Reed-Solomon Codes and a New Perspective on the Berlekamp-Massey Algorithm, *Symmetry*, 2019, Vol. 11 (1357). DOI: 10.3390/sym11111357.
9. Ceria M., Mora T., Sala M. Help: a sparse error locator polynomial for BCH codes, *Applicable Algebra in Engineering, Communication and Computing*, 2020, Vol. 31, pp. 215–233. DOI: <https://doi.org/10.1007/s00200-020-00427-x>.
10. Almuzakkia M. Z., Oharac K. Computing general error locator polynomial of 3-error-correcting BCH codes via syndrome varieties using minimal polynomial [Electronic resource], *ISCS, Selected Papers*, 2015, pp. 80–85. Access mode: https://mafiadoc.com/computing-general-error-locator-polynomial_5bad280f097c479e798b4727.html.
11. Freyman V. I. Research of the reed-solomon codes characteristic for realization within control systems devices, *Radio Electronics, Computer Science, Control*, 2019, Vol. 3, pp. 143–151. DOI: <https://doi.org/10.15588/1607-3274-2019-3-1>.
12. Liang Z., Zhang W. Efficient Berlekamp-Massey Algorithm and Architecture for Reed-Solomon Decoder, *Journal of Signal Processing Systems*, 2017, Vol. 86, Issue 1, pp. 51–65. DOI: <https://doi.org/10.1007/s11265-015-1094-1>.
13. Caruso F., Orsini E., Sala M., Tinnirello C. On the Shape of the General Error Locator Polynomial for Cyclic Codes // *IEEE Transactions on Information Theory*, 2017, Vol. 63, Issue 6, pp. 3641–3657. DOI: 10.1109/TIT.2017.2692213.
14. Bucerzan D., Dragoi V., Richmond T. The simple roots problem [Electronic resource], *Proceedings of the Romanian Academy, (Special issue), Cryptology Science*, 2017, Vol. 18, pp. 317–332. Access mode: <https://acad.ro/sectii2002/proceedings/doc2017-4s/03artSupl.pdf>.

Received 19.06.2020.

Accepted 20.08.2020.

МОДИФІКОВАНИЙ АЛГОРИТМ ПОШУКУ КОРЕНІВ ПОЛІНОМА ЛОКАТОРІВ ПОМИЛОК ПРИ ДЕКОДУВАННІ БЧХ КОДІВ

Крилова В. А. – канд. техн. наук, доцент кафедри автоматизації та управління в технічних системах Національного технічного університету «Харківський політехнічний інститут», Харків, Україна.

Тверитникова О. Є. – д-р наук, професор кафедри інформаційно-вимірювальних технологій та систем Національного технічного університету «Харківський політехнічний інститут», Харків, Україна.

Васильченко О. Г. – канд. техн. наук, доцент кафедри автоматизації та управління в технічних системах Національного технічного університету «Харківський політехнічний інститут», Харків, Україна.

Колісник Т. П. – канд. пед. наук, доцент кафедри інформаційних технологій та кібербезпеки Харківського національного університету внутрішніх справ, Харків, Україна.

АНОТАЦІЯ

Актуальність. У телекомунікаційних та інформаційних системах зв'язку з підвищеною шумовою складовою використовуються перешкодостійкі циклічні БЧХ та коди Ріда-Соломона. Коригування та виправлення помилок в повідомленні вимагає ефективних методів декодування. Одним з етапів процедури декодування РС і БЧХ кодів для визначення позицій спотворень є пошук коренів полінома локаторів помилок. Обчислення коренів многочлена, особливо у кодів зі значною коректне здатністю, є трудомісткою завданням, що вимагає високої обчислювальної складності. Тому удосконалення методів декодування БЧХ і РС кодів, що дозволяють зменшити складність обчислень, є актуальним завданням.

Мета роботи. Дослідження і синтез прискореного алгоритму пошуку коренів полінома локаторів помилок, представленого у вигляді афінного многочлена з коефіцієнтами в кінцевих полях, який дозволяє прискорити процес декодування БЧХ і РС кодів.

Метод. Класичний метод пошуку коренів на базі алгоритму Ченя виконується за допомогою арифметики кінцевих полів Галуа і трудомісткість розрахунків, в даному випадку, залежить від кількості операцій додавання і множення. Для лінеаризованих поліномів процедура пошуку коренів, заснована на двійковій арифметиці та здійснюється з урахуванням значень отриманих на попередніх етапах обчислення, що забезпечує мінімальне число арифметичних операцій.

Результати. Розроблено прискорений алгоритм обчислення значень полінома локаторів помилок у всіх точках кінцевого поля $GF(2^m)$ для лінеаризованих многочленів на базі методу Берлекемпа-Мессі. Алгоритм містить мінімальну кількість операцій додавань, за рахунок використання на кожному етапі обчислень, значень отриманих на попередньому кроці, а також виконання складання в кінцевому полі $GF(2)$. Запропоновано модифікований метод пошуку коренів для афінних поліномів над кінцевими полями, що дозволяє визначити позиції помилок в кодовому слові під час декодування циклічних БЧХ і РС кодів.

Висновки. Наукова новизна роботи полягає в удосконаленні алгоритму обчислення коренів многочлена локаторів помилок, коефіцієнти якого належать до елементів кінцевого поля. При цьому спрощується процедура декодування циклічних БЧХ і РС кодів, за рахунок зниження обчислювальної складності одного з етапів декодування – знаходження позицій помилок з використанням модифікованого алгоритму Берлекемпа-Мессі. Дані факти підтверджені результатами програмного моделювання алгоритму пошуку коренів полінома локаторів помилок. Показано, що застосування прискореного методу дозволяє досягти вигоди по швидкодії в 1,5 рази.

КЛЮЧОВІ СЛОВА: БЧХ коди, поліном локаторів помилок, пошук Ченя, алгоритм Берлекемпа-Мессі, коди Ріда-Соломона.

МОДИФИЦИРОВАННЫЙ АЛГОРИТМ ПОИСКА КОРНЕЙ ПОЛИНОМА ЛОКАТОРОВ ОШИБОК ПРИ ДЕКОДИРОВАНИИ БЧХ КОДОВ

Крылова В. А. – канд. техн. наук, доцент кафедры автоматизации и управления в технических системах Национального технического университета «Харьковский политехнический институт», Харьков, Украина.

Тверитникова Е. Е. – д-р наук, профессор кафедры информационно-измерительных технологий и систем Национального технического университета «Харьковский политехнический институт», Харьков, Украина.

Васильченко О. Г. – канд. техн. наук, доцент кафедры автоматизации и управления в технических системах Национального технического университета «Харьковский политехнический институт», Харьков, Украина.

Колесник Т. П. – канд. пед. наук, доцент кафедры информационных технологий и кибербезопасности Харьковского национального университета внутренних дел, Харьков, Украина.

АННОТАЦИЯ

Актуальность. В телекоммуникационных и информационных системах связи с повышенной шумовой составляющей используются помехоустойчивые циклические БЧХ и коды Ріда-Соломона. Корректировка и исправление ошибок в сообщении требует эффективных методов декодирования. Одним из этапов процедуры декодирования РС и БЧХ кодов для определения позиций искажений является поиск корней полинома локаторов ошибок. Вычисление корней многочлена, особенно у кодов со значительной корректирующей способностью, является трудоемкой задачей, требующей высокой вычислительной сложности. Поэтому усовершенствование методов декодирования БЧХ и РС кодов, позволяющих уменьшить сложность вычислений, является актуальной задачей.

Цель исследования. Исследование и синтез ускоренного алгоритма поиска корней полинома локаторов ошибок, представленного в виде аффинного многочлена с коэффициентами в конечных полях, который позволяет ускорить процесс декодирования БЧХ и РС кодов.

Метод. Классический метод поиска корней на базе алгоритма Ченя выполняется с помощью арифметики конечных полей Галуа и трудоемкость расчетов, в данном случае, зависит от количества операций сложения и умножения. Для линейризованных полиномов процедура поиска корней, основанная на двоичной арифметике, осуществляется с учетом значений полученных на предыдущих этапах вычисления, что обеспечивает минимальное число арифметических операций.

Результаты. Разработан ускоренный алгоритм вычисления значений полинома локаторов ошибок во всех точках конечного поля $GF(2^m)$ для линейризованных многочленов на базе метода Берлекэмп-Мессе. Алгоритм содержит минимальное число операций сложений, за счет использования на каждом этапе вычислений, значений полученных на предыдущем шаге, а также выполнения сложения в конечном поле $GF(2)$. Предложен модифицированный метод поиска корней для аффинных полиномов над конечными полями, позволяющий определить позиции ошибок в кодовом слове при декодировании циклических БЧХ и РС кодов.

Выводы. Научная новизна работы состоит в усовершенствовании алгоритма вычисления корней многочлена локаторов ошибок, коэффициенты которого принадлежат элементам конечного поля. При этом упрощается процедура декодирования циклических БЧХ и РС кодов, за счет снижения вычислительной сложности одного из этапов декодирования – нахождения позиций ошибок с использованием модифицированного алгоритма Берлекэмп-Мессе. Данные факты подтверждены результатами программного моделирования алгоритма поиска корней полинома локаторов ошибок. Показано, что применение ускоренного метода позволяет достичь выигрыша по быстродействию в 1,5 раза.

KEYWORDS: БЧХ коды, полином локаторов ошибок, поиск Ченя, алгоритм Берлекэмп-Мессе, коды Рида-Соломона.

ЛІТЕРАТУРА / LITERATURA

1. Berlekamp E. R. Algebraic Coding Theory / E. R. Berlekamp. – New York : McGraw-Hill, 1968. – 466 p.
2. Blahut Richard E. Algebraic Codes for Data Transmission / E. Richard Blahut. – Cambridge : Cambridge University Press, 2003. – 482 p.
3. Lin S. Error control coding: fundamentals and applications / S. Lin, D. J. Costello. Prentice-Hall Inc : Printed in the United States of America, 2004. – 624 p.
4. Truong T. K. Fast algorithm for computing the roots of error locator polynomials up to degree 11 in Reed-Solomon decoders / T. K. Truong, J. H. Jeng, I. S. Reed // IEEE Transactions on Communications. – 2001. – Vol. 49, Issue 5. – P. 779–783. DOI:10.1109/26.923801.
5. Nabipour S. Error Detection Mechanism Based on Bch Decoder and Root Finding of Polynomial Over Finite Fields / S. Nabipour, J. Javidan, F. Zare Gholamreza // Journal of Mathematics and Computer Science. – 2014, Issue 4. – P. 271–281. DOI: <http://dx.doi.org/10.22436/jmcs.012.04.03>.
6. Fedorenko S. V. Finding roots of polynomials over finite fields / S. V. Fedorenko, P. V. Trifonov // IEEE Transactions on Communications. – 2002. – Vol. 50, Issue 11. – P. 1709–1711. DOI: 10.1109/TCOMM.2002.805269.
7. Improved hybrid algorithm for finding roots of error-locator polynomials / [S. Fedorenko, P. Trifonov, E. Costa, H. Haas] // European Transactions on Telecommunications. – 2003. – Vol. 14, Issue 5. – P. 411–416. DOI: <https://doi.org/10.1002/ett.936>.
8. Bras-Amorós M. The Symmetric Key Equation for Reed-Solomon Codes and a New Perspective on the Berlekamp-Massey Algorithm / M. Bras-Amorós, E. Michael O’Sullivan // Symmetry – 2019. – Vol. 11 (1357). DOI: 10.3390/sym11111357.
9. Ceria M. Help: a sparse error locator polynomial for BCH codes / M. Ceria, T. Mora, M. Sala // Applicable Algebra in Engineering, Communication and Computing. – 2020. – Vol. 31. – P. 215–233. DOI: <https://doi.org/10.1007/s00200-020-00427-x>.
10. Almuzakkia M. Z. Computing general error locator polynomial of 3-error-correcting BCH codes via syndrome varieties using minimal polynomial [Electronic resource] / M. Z. Almuzakkia, K. Oharac // ISCS, Selected Papers. – 2015. – P. 80–85. – Access mode: https://mafiadoc.com/computing-general-error-locator-polynomial_5bad280f097c479e798b4727.html.
11. Фрейман В. И. Исследование характеристик кодов Рида-Соломона для реализации в устройствах систем управления / В. И. Фрейман // Радіоелектроніка, інформатика, управління. – 2019 (3). – С. 143–151. DOI: <https://doi.org/10.15588/1607-3274-2019-3-1>.
12. Liang Z. Efficient Berlekamp-Massey Algorithm and Architecture for Reed-Solomon Decoder / Z. Liang, W. Zhang // Journal of Signal Processing Systems. – 2017. – Vol. 86, Issue 1. – P. 51–65. DOI: <https://doi.org/10.1007/s11265-015-1094-1>.
13. On the Shape of the General Error Locator Polynomial for Cyclic Codes / F. Caruso, E. Orsini, M. Sala, C. Tinnirello // IEEE Transactions on Information Theory. – 2017. – Vol. 63, Issue 6. – P. 3641–3657. DOI: 10.1109/TIT.2017.2692213.
14. Buczerzan D. The simple roots problem [Электронный ресурс] / D. Buczerzan, V. Dragoi, T. Richmond // Proceedings of the Romanian Academy, (Special issue), Cryptology Science. – 2017. – Vol. 18. – P. 317–332. – Режим доступа: <https://acad.ro/sectii2002/proceedings/doc2017-4s/03artSupl.pdf>.

UDC 004.942:656.61.052

AUTOMATED IDENTIFICATION OF AN OPERATOR ANTICIPATION ON MARINE TRANSPORT

Nosov P. S. – PhD, Associate Professor of Navigation and electronic navigation systems Department, Kherson State Maritime Academy, Ukraine.

Popovych I. S. – Dr. Sc., Professor of the Department of General and Social Psychology, Kherson State University, Ukraine.

Cherniavskiy V. V. – Dr. Sc., Professor, Rector of Kherson State Maritime Academy, Ukraine.

Zinchenko S. M. – PhD, Senior Lecturer of Ship Handling Department, head of the laboratory of electronic simulators, Kherson State Maritime Academy, Ukraine.

Prokopchuk Y. A. – Dr. Sc., Associate Professor, Leading Researcher, Department of systems analysis and control problems, Institute of Technical Mechanics, National Academy of Sciences, Ukraine.

Makarchuk D. V. – PhD, Senior Lecturer of Ship Handling Department, Kherson State Maritime Academy, Ukraine.

ABSTRACT

Context. The article discusses approaches to anticipation identification being an essential part of the decision-making process done by the operator by using the example of a sea captain in ergatic systems of critical infrastructures in the sea transport management. The mentioned above aspect of anticipation of operators can be regarded as being a complex form of human-machine interaction and, certainly, claims for further elaboration of information and tools to be used.

Objective. The way to approach development is taken as being based on an information analysis of the full range of trajectories of decision-making by operators at the time of performing complex multi-stage actions. These items are rooting out of their adopted strategy of human-machine interaction. Besides, it leads to the formation of a metric being able to algorithmically represent the enormous number of variants. It can be done taking into account conditions of combinatorial representation in terms of the geometric theory of groups on the Cayley graph.

Method. Being a part of the approach elaboration the having been obtained during the analysis of the database of navigation simulators mathematical model of experimental data collecting and processing succeeded to be constructed. To confirm the formal-algorithmic approach a simulation was challenged to be carried out helping to form the trajectory of the operator's decision making in critical situations. It was felicitously performed basing on the three-factor ERO-AEA-EAPI model. Thus, the algebraic and software representation of the metric decision space is noticed to uncover approximate complex human-machine interactions in uncertain environments. As a result, the converting process of data of the main subject of critical infrastructure (i.e. the operator) into knowledge is able to be coped with. In addition, factors possible to be gauged in the proposed metric are able to be uncovered.

Results. In order to carry out the feasibility assessment of the developed approach as well as formal-algorithmic ones, an experiment was performed by using the Navi Trainer 5000 navigation simulator (NTPRO 5000). During having one of the most troublesome operations i.e. mooring we wanted the server data to be analyzed. As a result, data about anticipation being shaped as triangular constructs in the quasi-isometric space of Cayley graph is reported to have been obtained. The automated neural networks being used for result obtaining led to delivering of the possibility to get multiple data regression and to analyze the relationships of many independent variables. It is considered to be clear evidence due to having found out results of scattering and reliability diagrams.

Conclusions. The having been presented in the investigations formal-algorithmic approach together with the developed software tools and the approaches of converting data into knowledge about operator anticipation are said to embrace the possibility to classify and to identify individual decision-making strategies when managing a vessel and to predict the likelihood of poor consequences. With regarding to the cogency of the proposed approach and models these issues happen to have been successfully justified by means of the automated processing of experimental data.

KEYWORDS: decision support systems; knowledge identification systems; operator model; computer navigation simulators; management of risks; human factor analysis; automated control systems.

ABBREVIATIONS

DSS is a decision support system;
NTPRO 5000 is a navigation simulator “Navi Trainer 5000”;
ARPA is an automatic radar plotting aid;
AIS is an Automatic Identification System;
ECDIS is an Electronic Chart Display and Information System;
RADAR is a radar station;
ERO is an expected results of operations;
AEA is an anticipated event awareness;
EAPI is an expected attitude towards participants in interpersonal interaction.

NOMENCLATURE

X_g, Y_g is a metric space;
 d_{x_g}, d_{y_g} is a space metric;
 δ is a geodesic neighborhood;
 x' is a signs of anticipation;
 $(X_t^*)\alpha$ is an influence of new data;
 W_i is a conditions;
 Z is a mooring task;
 a_t is a detalization;
 b_t is a generalization;
 f is a quasimetric attachment;

c_g, k are constants;
 a is a maneuvering with a change in the operating mode of the main engine;
 b is a shift pen rudder;
 c is a change of the operating mode of the thruster;
 p is a point located on the geodesic segment;
 c_a is a bounded quasi-isometric mapping;
 x_g, y_g, z_g are vertices;
 γ is a geodesic segment (path);
 r is a geodesic segment length;
 S_i is an area of the flat;
 (M, d) is a metric space;
 a_g, b_g are an any two points on sides;
 x is a situation;
 $B_t(x)$ is an action plan (detail);
 $C_t(x)$ is a summarizing data;
 $\{Ind\}$ is a forms inductors;
 D_j is an information domains;
 $P(x, z)$ is a predicate;
 $A(x)$ is a set of probable solutions;
 Ω_t is a set of rules is primarily formed;
 H_t is a data generation conditions;
 $\{S^*\} \alpha$ is a favorable treatment;
 U_t is a possibility of having options;
 e is an activity of captain;
 V is an objective reality, regularity;
 EZ is an energy and motivation of the captain;
 τ is a basis of the tests;
 $\{\Xi\}$ is a destined for observing the constructing domain digraphs experience;
 θG is a set of elementary operations;
 $\{Ag\}$ is a micro social environment limited.

INTRODUCTION

It goes without saying that the processes involving human-machine interaction in ergatic systems are sure to have always been a crucial field directly affecting the final result. In the contemporary critical infrastructures with an operator having in the information circuit, under the great security influence these processes are to be drawn attention to as reflecting great interest [1]. Modern information technologies demonstrating competencies in finding solutions aimed at ensuring the safety of ergatic systems are using the example of maritime transport to a small degree. The human factor manifestations are named to have always been main contributors to a significant increase of the likelihood of faulty situations.

Signaling is usually made in the form of a warning directive to the vessel crossing dangerous areas or approaching target vessels (Fig. 1), obstacles or dangerous

isobaths so, therefore, a safety circuit is mandatory to be used (Fig. 2.) [2].



Figure 1 – Identification of hazardous areas of navigation

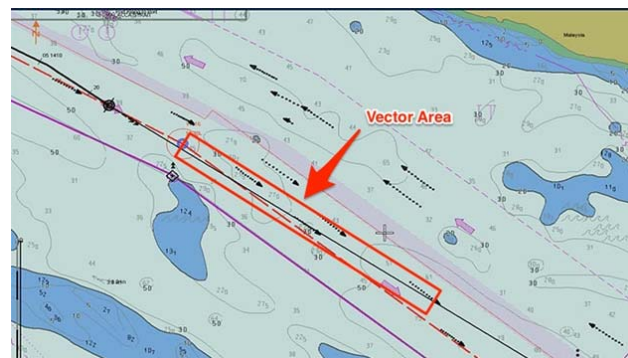


Figure 2 – Display of the ship safety frame

The operator himself is able to choose and even create his own circuits for having being alerted by the security system experience. The ECDIS system is noticed to be having enough tools for automated security settings by using electronic navigation charts. In addition, other troubles to be mentioned happening like signaling due to having problems in the engine room, rough weather conditions can be successfully coped with. At the same time, on the opposite, there is little evidences speaking about tools that are involved into taking into account the operator's response to the data stream and signals of navigation devices. Selecting highly qualified operators might be a key issue in the attempt to solve this problem but the complication to be paid special attention to is not being aware of the having been used shaping experience [3]. Besides, when constructing automated systems frame aiming to improve security level in this meaning, it is important to be focused on the "average" operator experience and qualification. Nowadays they tend to be more than 70% of the total. Previously having been implemented reviews made us face the challenge of highlighting the questions of having negative influence of individual groups of factors on the operator's behavior pattern, concentration of attention and communication features in the group [4–6].

However, questions related to the planning of the trajectory of actions and the expectation of a future result due to their individual formation pattern are still remained to be not enough discussed. Certainly, in order to obtain results in this field it is considered to be complicated to determine the approaches for identifying information signs and signals. Nevertheless, exactly these issues make

it possible to build automated means for improving safety in maritime transport.

Thus, the problem is highlighted reflecting having the negative manifestation of operator anticipation as a component of the human factor. Analysis of accidents as well as work with cadets during training drills at navigation simulators at the Kherson State Maritime Academy (Ukraine) proves the established hypothesis of anticipation being decisive cornerstone in making decisions correct beyond any doubt. This pattern of human factor behavior especially manifests itself in the following conditions: an operator has already had experience, for example, captain or his first assistant, or he navigates in the water area he is quite get acquainted with. It must be emphasized that avoiding the negative manifestation of anticipation is highly likely to happen by automated means for its identification and manifestations in dynamics relative to each operator.

Hence, **the object of research** frames the question of the process of automated identification of anticipation of operators while performing complex maneuvers in the management of sea transport.

The subject of the study is models and algorithms that implement the process of automatic identification of operator anticipation in critical and difficult-to be predicted situations.

The purpose of the study is taken to be to introduce and promote a mathematical and algorithmic justification for the development of a module of a DSS in order to identify the beginning point of operator's anticipation and its dynamics in the presented metric to meet or to improve safety level in maritime transport.

The purpose of the article is chosen to be the applying special purpose solution method embracing the following problems:

1. To specify a formal approach for the formation of a functionally defined space Y_g with a metric d_{y_g} , to display the manifestations of anticipation of the operator.

2. To approximate the decision-making process made by the operator basing on the three-factor model ERO-AEA-EAPI in the form of a geometrically defined trajectory on the surface of Cayley graph X_g with a metric d_{x_g} , with an intention to synchronize the data of the Transas navigation simulator NTPRO 5000 and the encoded trajectory.

3. To determine extrema on the Cayley graph able to convert the surface to quasi-isometric item and allowing to obtain the distortion value in the form δ of a-neighborhood on the geodesic triangle ERO-AEA-EAPI. According to carrying out experience the operator's anticipation centers seem to be activated and seen.

4. To conduct an experiment using the Transas navigation simulator NTPRO 5000. It would be involved into having the possibility of automatically manner build the triangular space of the ERO-AEA-EAPI based on the results of the navigation tasks and identify the centers, magnitude and dynamics of operators' anticipation.

5. To formally introduce an information model of the formation of operator anticipation and identify signs of anticipation x' , as well as the influence of new data $(X_t^*)\alpha$ on its manifestation in the form of a thin section X_{t+1}^* under the conditions W_t, Z and psycho-physiological parameters of the operator.

6. To develop a model of the dynamics of the centers of anticipation of operators in the form of a Shukhart-Deming spiral cycle from detalization a_t to generalization b_t and back a_{t+1} . This issue will be able to provide the possibility of having negative manifestations of anticipation predicted in a discrete form.

7. To construct a multiple regression using the experimental data of the trajectories of decision-making by operators by means of automated neural networks. Further actions proposing to be taken include their training process to be involved with an aim to organize the effective work of DSS and, as a result, to identify the negative manifestation of anticipation in maritime transport operators.

Hence, the best tackle of current spoken above challenges means dealing with getting the opportunity to succeed in achieving problem solution of identification and prediction of anticipation of operators in maritime transport. The allocated experimental data is intended to deliver further probable development of DSS with a goal of creating and enhancing more sustainable safety environment in maritime transport by diminishing the negative manifestations of the human factor.

1 PROBLEM STATEMENT

The research dedicated to the operator mental states map designing in rough environment reflected the idea of having identified points on its surface mattered. It must be noted that each operator is about to have this surface functioning in up-to individual mode due to its basing on personal preferences and behavior patterns as well as on a complex system of social expectations. From the first point of view, the base points of the surface are not noticed to be connected with each other. However, further reconsideration can propose getting the opportunity of building connections by selecting construction elements using terminology of the theory of geometric groups [7].

Thereby, construction elements form segments and the last ones tend to build a geometric structure i.e. functionally defined space. Consequently, the points connected in space form a metric by the help of the given distances between them. Taking into consideration the fact of having all points of the obtained space Y_g with metric d_{y_g} , included in the primary Cayley graph X_g with metric d_{x_g} , the conclusion worth taking is that this metric is induced. So, these points will certainly limit the primary Cayley graph, convert it to the induced pseudo-metric where the points are connected by geodesic segments.

It is vividly seen that for the particular person mental state of being under the influence of his social expectation

factors geometric structures will be shaped as a quasimetric attachment f in the form of $f: X_g \rightarrow Y_g$, if

$$d_{x_g} \quad d_{y_g}$$

$\exists k, c \quad \forall x_g, y_g \in X_g$. To meet this issue requirements for the validity that the distance between the points of the spaces X_g and Y_g can only change linearly must be observed. This condition delivers the opportunity to save the metric of the primary Cayley graph:

$$\begin{aligned} d_{y_g}(F(x_g), F(y_g)) &\dots \Rightarrow d_{y_g}(f(x_g), f(y_g)) \leq \\ &\leq k d_{x_g}(x_g, y_g) + c_g \Rightarrow \frac{d_{x_g}(x_g, y_g)}{k} - c_g \leq \\ &\leq d_{y_g}(f(x_g), f(y_g)) \leq k d_{x_g}(x_g, y_g) + c_g. \end{aligned}$$

Moreover, the nature of this dependence is quasi-isometric one because the distance between the points $f(x_g)$ and $f(y_g)$ varies linearly depending on the indicated above constants c_g and k .

Suppose the application of the proposed approach to the formation of space of social expectations of the individual basing on the results of the research made by Professor I. Popovych at al. [8]. His inquiry suggests a three-factor model to be used consisting of the following areas:

a) ERO; b) AEA; c) EAPI.

Being guided by the mentioned above options, the conclusion can be drawn that the personality while being involved into carrying out activities is basing mostly on three strategies causing definite deeds. Thereby, his social expectations might be founded on these three parallel directions and superimposed on the trajectory of action.

2 LITERATURE REVIEW

The research problems are rooted out from the contradictions existing between the needs for automated identification of operator behavior motives in ergatic systems and the capabilities of navigation information systems [9]. In having been studied inquiries related to these issues scientific researches tend to get risk assessment by mathematical processing of the parameters of the information model of the vessel and its dynamics in the water area [10–12]. Others are focused mainly on the enhancement of information tools for the analysis of physiological characteristics of sea transport operators [13, 14]. At the same time, the decision support systems questions can not be avoided as being essential in facilitating safety marine environment reducing the human factor of the operator by analyzing the decisions made during holding human-machine interaction [15–18]. Furthermore, the task of identifying the manifestations of the human factor itself can be found as being much more problematic one in comparison with ordinary work performed with signals and data from navigation devices RADAR, ARPA, AIS, ECDIS, etc. In the meantime, the deliberate relevance of such investigation is noticed to be expending greatly thanks to automated systems being involved. To some extent, this practice let them respond to manifestation of

the negative human factor by automatical diverging in the principles of E-navigation [19, 20].

Contemporary approaches to the sustainable development of automated and intelligent systems are more and more involved into facing challenges to respond the need for undertaking thorough analysis of operator action planning [21, 22].

Nevertheless, well-known widely-spread literature is unable to deliver references to efficient automated systems capable of carrying out identification of operator anticipation manifestations as an indicator of the strategy of decisions made in ergatic systems of maritime transport.

3 MATERIALS AND METHODS

There is the hypothesis that due to the psychological characteristics of the individual at each interval (stage) of the trajectory of his activity one factor-direction is considered to act as a dominant one, being dominant. Moreover, each fragmented trajectory of human activity has three dominants in each direction factor, thereby, forming a triangular structure having non-trivial individual-personality properties. They turn out to have scientific and practical significance for the development of the methodology of social expectations of the individual.

For the sake of confirming this hypothesis a simulation of the trajectory of a person's behavior in critical situations is supposed to be performed. One peculiarity is to be made is that independent decision making is required to be based on the three-factor model of ERO-AEA-EAPI.

Suppose, as a modeling environment critical infrastructure (for example, maritime transport) would be chosen to be spoken about and as an operator, a watch assistant of the captain or captain. Besides, imagine, the environment where the vessel will be operating and maneuvering is said to be one of the most complicated ones (coming up to the pier and mooring operation). To completely satisfy and meet the requirements such options as experience and professional skills of senior assistant of a captain or other higher positions must be taken into consideration.

When constructing the Cayley graph the three axes a , b , c would be proposed to be used basing on the reasons that when mooring combinations of control actions rests on three main types of maneuvers:

a – maneuvering with a change in the operating mode of the main engine;

b – shift pen rudder;

c – change of the operating mode of the thruster.

As a rule, taking into account huge number of external and internal factors influencing on, the captain gets used to being guided by combinations of basic maneuvers monitoring the situation in real time mode. Gained experience and mind-motor skills are able to be highly likely the most influential factors on choosing the particular strategy. These items which brings an additional effect that polar influences the final result of the maneuver.

Thus, to identify the path of the skipper’s actions a metric space is proposed to be constructed. It would definitely take into account the entire spectrum of options for the those actions in the context of a three-factor model in the form of a Cayley graph basing on the geometric theory of groups.

It must be underlined that this graph, according to the axiomatics of the formal approach, is noticed to be starting from the base point e and to be spreading in three axes. The metric of the graph does not change with its spreading’s, however, for the image to be observed compactly, each level of construction is scaled a multiple $\frac{2}{3}$. So, the Cayley graph obtained in this way could be depicted as following (Fig. 3):

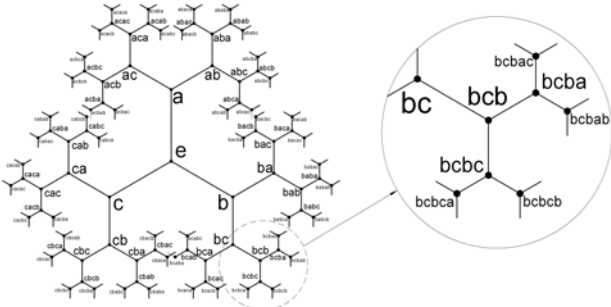


Figure 3 – Cayley’s graph of the actions of the skipper in the three-factor model

As it can easily be seen from Fig. 3, the Cayley graph is not observed changing structurally. To some extent at any point the principle of the geometric distribution of edges can be functioning in the same way. This tendency confirms the appropriateness of geometric group theory sampling techniques for further development of this study. In order to analyze this approach, an experiment was carried out using the Transas navigation simulator NTPRO 5000 at the Kherson State Maritime Academy (Ukraine) (Fig. 4).

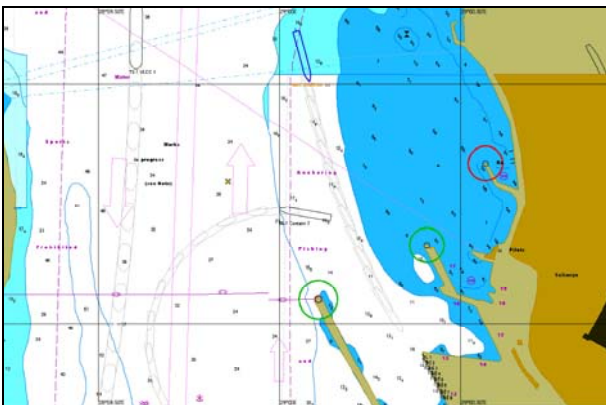


Figure 4 – A fragment of the operation on the NTPRO 5000 simulator

Before starting of the experiment students, navigators-to-be, had been pre-screened for upcoming behavior pattern with the following analysis of the actions letting us

get identified all preferences regarding directional factors during the mooring operation.

In order to improve the experiment quality and leveling random factors a mooring operation is reported to have been performed in three stages. In each of the stages an extremum was determined by three directional factors towards particular captain. Consequently, three trajectories were obtained mapped onto fragments of the Cayley graph (Fig. 5).

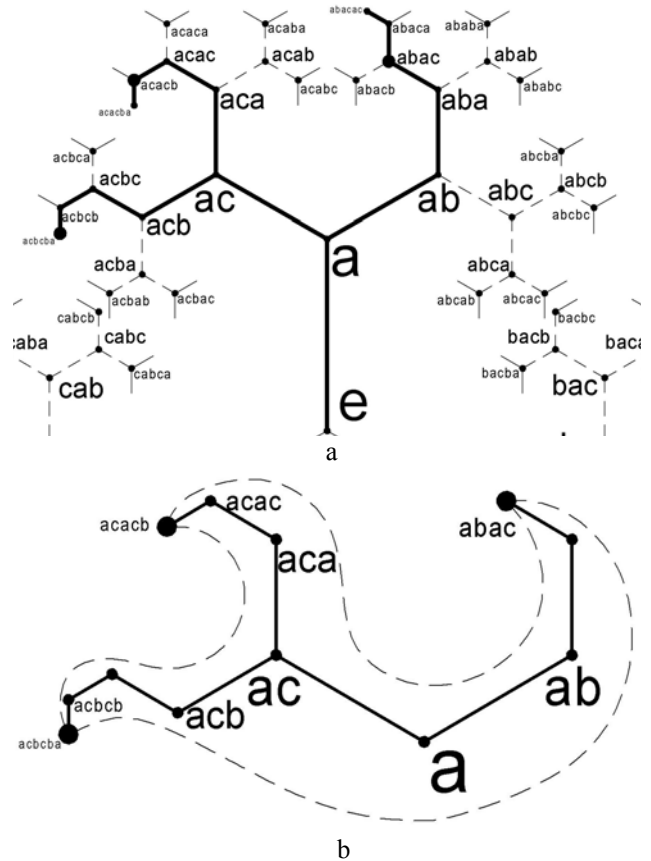


Figure 5 – The set of trajectories on the Cayley graph

Fig. 5a shows that all three trajectories begin with a factor-direction $a \in A$, indicating the fact of navigator’s having already been formed by the carrying out moment stages action plan. To a certain point, this is real truth due to the fact that the skipper is experienced and has well-established behavior algorithms in cases typical to this situation. Having highlighted and connected the extremum points ($acbca$, $acacb$, $abac$) we get a geometric figure forming a triangle. Attention must be paid to the evidence of functioning each extremum point as a determiner of the length of the trajectories of different sizes (usually from 4 to 6 stages) (Fig. 5 b).

The resulting structure is defined by the central $a-ac$, geodesic segment limiting the outgoing trajectories within a constant c_a , such that (1):

$$f(x_g) \exists c_a : \forall a \in A, \exists p \in F(X_g) : d_a(a, p) \leq c_a. \quad (1)$$

Thusly, this structure is not said to be anything more than c_a – bounded quasi-isometric mapping of the image describing the metric of social expectations of the skipper in a set of typical situations. The introduced constants are not noticed to distort the essence and linear dependence of the three-factor model of the ERO-AEA-EAPI. This data deliver us the possibility of getting generalization the observed properties of the formation of behavior trajectories. The obtained for each trajectory data is fragmented and displayed in Table 1 and according to it, the main extremes turn out to be analyzed and, as a result, it is individual resurfacing of the navigator’s social expectations is highly likely to happen (Fig. 6).

Table 1 – The data of the trajectories of the geometric structure

№	Trajectory	Trajectory stages					
		a_1	ac_1	acb	$acbc$	$acbcba$	$acbcba$
1	$acbcba$	a_1	ac_1	acb	$acbc$	$acbcba$	$acbcba$
2	$acacba$	a_2	ac_2	aca	$acac$	$acacba$	$acacba$
3	$abacac$	a_3	ab	aba	$abac$	$abacac$	$abacac$

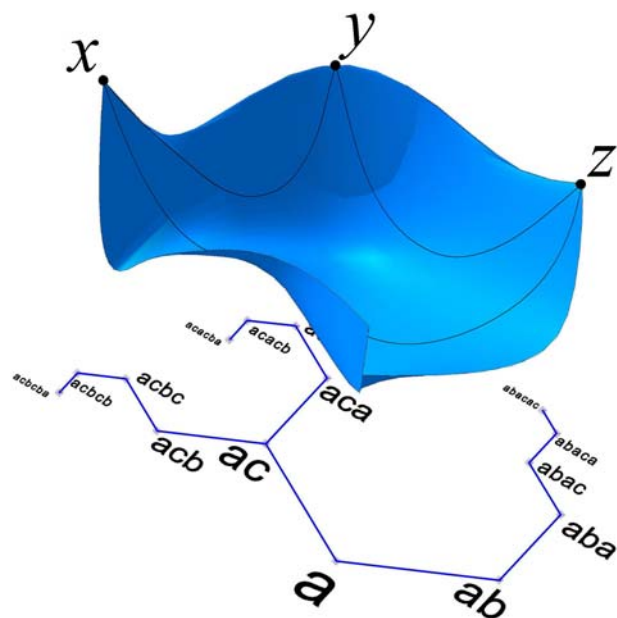


Figure 6 – The curved surface of the social expectations of the individual

Further observation of the specifics of the manifestations of factor directions makes us notice our moving away from the straightness of the Cayley graph due to the scatter of the values determined by its points.

The image of the graph itself can be regarded as an image on a curved surface formed by a matrix of values when forming the trajectories of the real actions of the skipper (Fig. 6). To a certain point, the formed surface is said to be an individual imprint of his social expectations in the situation under consideration (2):

$$(M, d) : \forall x_g, y_g, z_g \in M, \exists \gamma : [0, r] \rightarrow M. \quad (2)$$

The curvature determination of this surface contributes to providing the representation of the impact degree

of dominant factors on the situation and to displaying of individual habitual preferences being performed in a typical situation. As a result, possible future behavior pattern of the skipper is highly likely to be predicted. However, taking into account the complexity of this process, an indirect feature is proposed to be used due to the fact of getting real reflection of the deviation from the average indicators of the points. This indirect issue is noticed to be based on the ratio of the area formed by three points of the flat and the actual obtained surface.

The comparative analysis of the area of the flat $S_1 = 4210.04 \text{ mm}^2$ and the surface area $S_2 = 9531.18 \text{ mm}^2$ is equal to 2.263. Thus, a surge of factors regarding the situation turns out to have double time exceeding average experience which is considered to be treated as true indicator of skipper’s mental state.

Therefore, the goal points of social expectations are able to be determined for all three factors-directions of the model of the ERO-AEA-EAPI.

So, the formed space is δ hyperbolic [23] because (3):

$$\forall x_g y_g z_g \in M, \forall [x_g, y_g], [y_g, z_g], [x_g, z_g] \rightarrow M\delta. \quad (3)$$

Being proved by the fact that all the geodesic segments in our case are located in the same neighborhood surroundings common trajectory fragments are vividly seen. This issue is confirmed by transition properties (4):

$$\begin{aligned} \exists \delta : \forall \text{geodesic } \Delta \rightarrow \delta \text{ narrow} &\Rightarrow \exists \delta : \forall \text{geodesic } \Delta \rightarrow \\ &\rightarrow \text{diam}\{\text{center points } -a, b, c\} \leq \delta. \end{aligned} \quad (4)$$

It can be observed from formulas (3) – (5) that this way formed geometric structures have a central space being limited by points $(c_{x_g}, c_{y_g}, c_{z_g})$, so called δ neighborhood. In our case it will have a diameter equal to the edge $a-ac$ being completely appropriate for the c -limitation condition mentioned above (Fig. 7).

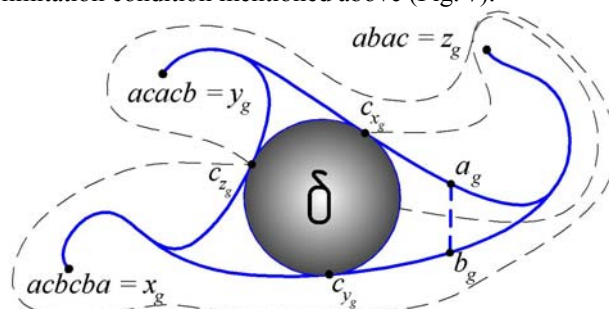


Figure 7 – Is δ -neighborhood surrounding on the geodesic triangle ERO-AEA-EAPI.

Depicted in Fig. 7 a geodesic triangle is quasimetrically embedded being a fragment of the Cayley graph into a hyperbolic space as its δ -finess, $[x_g, y_g] \cup [x_g, z_g] \cup [y_g, z_g]$ in the metric space (M, d) . This situation can easily be determined by the following condi-

tion: for the vertex z and for any two points a_g and b_g on its sides $[z_g, y_g]$ and $[z_g, x_g]$, the next coming item is observed:

$$d(z_g, a_g) = d(z_g, b_g) \leq \frac{1}{2}(d(z, y) + d(z, x) - d(y, z)) \Rightarrow \\ \Rightarrow d(a_g, b_g) \leq \delta, \delta > 0.$$

Consequently, that presented in the scope of this study three-factor model of ERO-AEA-EAPI on this Cayley graph group happens to be mentioned a hyperbolic one basing on the definitions of Gromov-Rips [24, 25]. From the spoken above the conclusion can be made about existing δ -neighborhood surroundings for a finite system of generators where the union of any two sides occur (Fig. 7). As it can be clearly seen, the curvature of the space ERO-AEA-EAPI seems to be situational one and depends on its generators. The result of this as follows: the problem of revealing the pattern of its formation as well as the complication of making classification of the navigator's mental states regarding to extreme points happen to surface.

4 EXPERIMENTS

In the light of the foregoing above objective an experiment was carried out using the Transas navigation simulator NTPRO 5000. This experience contributed into succeeding in having the space of the ERO-AEA-EAPI built basing on the results of the navigation tasks of three captains.

During the experiment double having been passing of the location by each of the captains in two similar situations in Bosphorus and Hong Kong was performed. In each case extreme points were identified showing that the obtained data is close enough in meaning, in average comprising about 2–5%. To a certain point, this issue indicates the awareness in performed decision-making process based on many years of experience. Despite the probable static nature of the identified indicators of the model of the ERO-AEA-EAPI, dynamics tended to be found in the form of the displacement of triangular geometric structures towards the geodetic space above the Cayley graph (Fig. 8). For example, being traced in two cases dynamics has various transformations. It must be emphasized that in the first case a significant expansion of the surface area formed by triangles by 65% is possible to be observed. This fact indicates that the captain No. 2 tends to be shifting the focus of the ERO-AEA-EAPI intentionally at the second stage.

So, for the sake of getting the extremum points figured out an automated system was developed allowing you to visually determine the surface of the triangulation (Fig. 8). This item is sure to indicate the balance of the ERO-AEA-EAPI in the mooring task Z .

It must be noted that the long-drawn-out experiment is reported to have been performed in the period of 3.5 years

with dozens of trajectories analyzed. Subsequently, the tendency was noticed to follow of having some cases with the dynamics of the displacement of the points of the extrema in time slices.

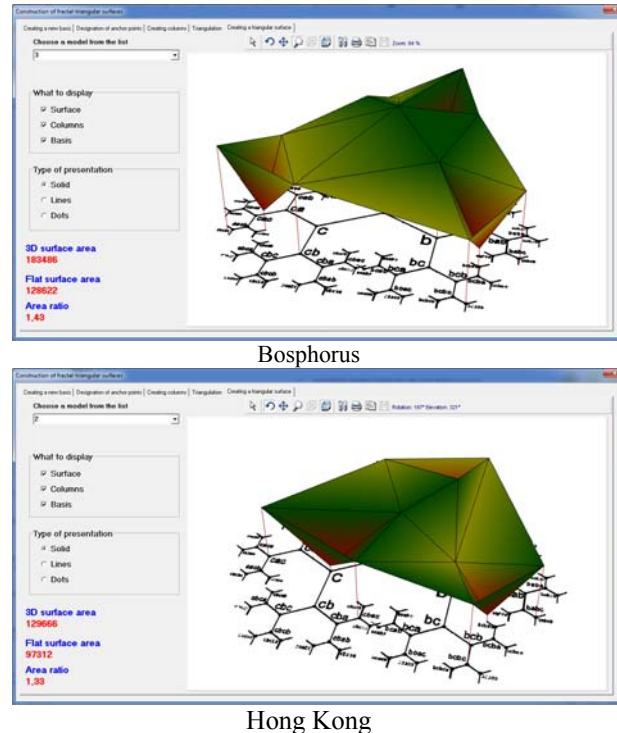


Figure 8 – Triangulation of the hyperbolic decision surface

The implemented research revealed a bias dependence depending both on the age of the captains and on their psychological states (Fig. 9).

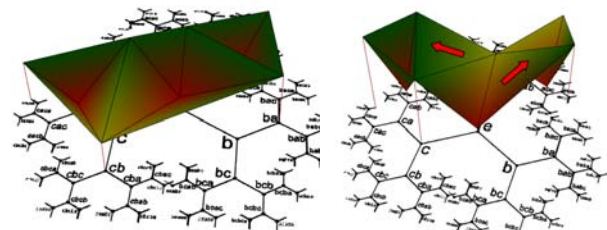


Figure 9 – Dynamics of the transformation of a hyperbolic surface

As a result, a hypothesis is highly likely to be done is that such actions are thought to be habitual characteristics of individuals being still on their way of development and improvement. They are sure to be setting new goals and striving to boost their skills and abilities associated with running the vessel and navigation. The only problem to be mentioned is our being uncertain about this. Appropriately to the aforementioned, there is a need in applying mechanism for the condition determinations for the aim of stimulating the captain to change the strategic trajectory of action.

In this regard, let us consider this situation in the formal theory of creativity framework [26]. According to it, the notion to be used is mooring Z , as a decision-making

task it is to be based on the captain's initial information about the situation $x \in X$, in the form of which the initial action plan is shaped $z \in Z$. The captain's early experience occupies a significant place at this very stage. For instance, with a sufficiently low functional entropy of the captain, the action plan can possibly be determined with a high degree of accuracy z/Z , leading to the enormous specification of the action plan $B_t(x)$ proposing the only probable- to be carried out-plan X_t^\wedge . Otherwise, the captain tends to be relying indirectly on experience of such type of tasks Z , summarizing data $C_t(x)$ while having $x \in X_t$. This experience seem to form $X_t \rightarrow X_{t+1}$ and depict $C_t : X_t \rightarrow 2X_t$ and $B_t : X_t \rightarrow 2X_t$.

Thus, the expansion of the spectrum of actions by the level of generality $C_t(x)$ forms inductors $\{Ind\}Z$, with regard to information domains D_j , affecting the focus shift of the ERO-AEA-EAPI and determining model $D_1 \{Ind\} D_1 \rightarrow D_2 \{Ind\} D_2 \rightarrow \dots \rightarrow D_j \{Ind\} D_j$.

As a result, the assumption seem to appear that focusing shift is possible to be surfaced when there is either entropy (unconscious action) or (conscious action) if there is a need in having one's social condition improved and, hence, in anticipation itself.

In addition it should be taken into account that detailing and generalization are being involved into complex way functioning: $\forall x \in X_t C_t(x) \cap B_t(x) = x$ is on the problem of mooring Z , while new social states – knowledge base is being formed and equal to: $x' \in C_t(x) \Leftrightarrow x \in B_t(x')$. However, these conditions can be reported as being valid only when non-random activity of captain e , case is observed such as: $\forall x \in X_t e(x) \rightarrow e(C_t(x))$.

Basing on the fact that all captains due to their qualification base a-priori act consciously rely on their prior experience and skills the mooring problem is to be determined by a predicate $P(x, z)$, such as $\forall x \in X \exists z \in Z P(x, z)$. While having it, the power of the set of probable solutions will be: $\forall x \in X z \in A(x) P(x, z) \ \& \ \forall x \in X z \notin A(x) \neg P(x, z)$, where $A : X \rightarrow 2Z$. In tough close to-emergency situations the captain is obliged to operate with data with a maximum accuracy a criterion which is said to be $\Omega_t, H_t = \{x | x \in X_t, B_t(x) \equiv x\}$. This issue enormously reduces and limits the possibility of having options $U_t = \cap x \in H_t C_t(x)$, with stimulating the extinction of the extreme points of the ERO-AEA-EAPI.

The manifestation of such situations are causing the captain's mechanical and unemotional actions in the cases when: $\exists x, x' \in H_t$ that $C_t(x) \cap C_t(x') = \emptyset, U_t = \emptyset$ and

$\forall x \in X_t^\wedge \exists! z x P(x, z)$, when such tough situation is highly likely to happen: $\forall x' \in C_t(x) \cap X_t^\wedge P(x', z)$.

Therefore, for mooring operation to succeed a trajectory, $\{x_n\} x \equiv \{x = x_0 < x_1 < \dots < x_n\} \subset X_t^\wedge$ is to be provided escaping from getting the entry to the close-to-emergency situation described by the set of $X_t^* = \{x^* | x^* \in X_t \exists! z \in Z P(x^*, z)\}$ or emergency one like $X_t^+ = \{x | x \in X_t P(x, u)\} = X_t \setminus X_t^\wedge$. It is worth adding that such situation may condition the time frame to be determined forming the pre-emergency state of the mooring operation $X_t = X_t^\wedge \setminus X_t^*$. However, the complexity of having it done is the fact in each captains' having individual temporary structures according to their background nature displayed in the trajectories of behavior. It makes ground to welcome necessity of determination an individually-oriented model of the captain taking into account the time phases of effectiveness of safe providing model of navigation.

On account of all spoken above, let us turn to the formal concept defining individual temporary "codes" of captains within the framework of tackling such indicated problem mooring operation of the vessel: $\forall x \in X_t^\wedge \exists! z \in Z P(x, z)$.

Premising, all three time phases of its solution are considered as being $X_t^\wedge = \cup z \in Z (X_t^\wedge)_z$; $X_t^* = \cup z \in Z (X_t^*)_z$; $\forall z \in Z (X_t^\wedge)_z \rightarrow e(X_t^*)_z$, with the transition to is undiscrete they staged by means of adding certain conditions, $V = (x \rightarrow zx)$ in individual way for each captain. Then, each piece of information $x \in X_t^\wedge$, affecting the decision-making strategy needs to be tracked.

It is worth mentioning that before the mooring, at a certain point in time, the set of rules is primarily formed $\Omega_t = \{\alpha = \langle x\alpha, z\alpha \rangle | x\alpha \in H_t\}_t$ and is unable to provoke spontaneous phase transitions. However, the surfacing the new ones, being quite unpredictable $(X_t^*)\alpha$, can cause leveling the preliminary plan and form an initial set of decision-making rules opposite to the first one $\{S^*\}\alpha = \{x^* \rightarrow z\alpha | x^* \in (X_t^*)\alpha\}$.

Experimental data provide clear evidence that the captain, due to lack of time, as a rule, is not used to change the initial plan not observing any threat to the vessel safety. At the same time, if there is time for the strategy to be changed W_t, Z is a cause for the coming out experience of a thin section X_{t+1}^* , anticipation. It is no doubt that it is able to produce both a positive effect in the captain's knowledge base and a negative one reducing the safety of navigation up to having catastrophic consequences. Con-

sequently, when the captain has enough time to form a chain of events of the form it looks like:

$$\begin{aligned} \Omega_t &= \{V\}_0 \rightarrow W_{t,Z}(\{V\}_0) \rightarrow W_{t,Z}^{\circ n}(\{V\}_0) \rightarrow \dots \\ &\rightarrow W_{t,Z}^{\circ n}(\{V\}_0) = \{S^*\}_t, \text{ Full}, \{V\}_t, \text{ Full} = \cup k = \\ &= 0, 1, 2, \dots, n \quad W_{t,Z}^{\circ k}(\{V\}_0), EZ \geq 0, \text{ it happens } H_t \rightarrow \\ &\rightarrow W_{t,Z}(H_t) \rightarrow W_{t,Z}^{\circ n}(H_t) = X_t^*, EZ \geq 0. \end{aligned}$$

5 RESULTS

In this way, we are to emphasize that the conditions are nominally highlighted. They are said to be affecting the captain's ability to reach a decision possible to introduce changes into his model of knowledge base and welcome new ideas regarding the current task $W_{t,Z}^{\circ k}(\{V\}_0)$. During the experiments two vectors of the captain's transition to a qualitatively new level of perception of the navigation situation were noticed to be identified: the first is a change of strategy during the implementation of the initial navigation plan (17%); the second one concerns the formation of a new preliminary plan (36%) after n homogeneous ones X_{t-m} . The remaining 47% percent of cases reflect not having changed of the strategy approach of the captains while performing the mooring operation even in spite of getting new information signals come out. Thus, the conclusion highly likely to be done is that captains capable of transition X_{t+1}^* , may successfully manage with differentiation between a situation with a familiar dataset and a situation with a new one. Simultaneously, the transition X_{t+1}^* is possible to be done through a spiral cycle from detailing a_t to generalizing b_t and back to detailing a_{t+1} .

So, generally this algorithm can beneficially be represented in the framework of the Shukhart – Deming cycle [28] completely confirming the research hypothesis

(Fig. 9, 10). The diagram catches that the cycle $a_t \rightarrow b_t \rightarrow a_{t+1}$ is a key issue in having shift process of the extreme points ERO-AEA-EAPI. Therefore, the mechanism for transition formation in the cycle must be determined by means of selecting individual tests for each captain.

It is worth noticing that provided data granted us the opportunity to determine the factors of permitted -to be done activity of the captains which is reported to become the basis of the tests τ when performing a wide range of navigation tasks – τ/T .

The formal presentation of the mechanism is considered to be defined as a digraph $G(\tau) = \{T \rightarrow e T'\}_\tau$. This case must oblige us to draw special attention to the identification of attention areas focused on primary details:

$$\forall \tau/T \{e_{t+1}(\tau/T)\} = E \tau/T, t(Z(t), \{G(a)\}, E).$$

The fact of the coherence $C(\tau/T)$ of the navigation operations during activation of τ and sufficient energy and time can be treated in such way as: if $e(\tau/T) = 1 \Rightarrow e(C(\tau/T)) = 1 \& e(T \setminus \tau) = 0$, while having it the shift is $\forall T \rightarrow T' \{e_{t+1}(T \rightarrow T')\} = ET \rightarrow T', t(Z(t), \{G(a)\}, E)$. So, basing on the principles of quantitative-qualitative transition we are to observe $Gv(\tau) \Rightarrow \exists G\Xi(\tau)$. The assumption that $G_{l+1}(\tau) = \theta G, \tau(G_l(\tau))$, where $l=1, 2, \dots, m$, and $\theta G, \tau$ is a set of elementary operations of digraph development is proposed to be paid attention.

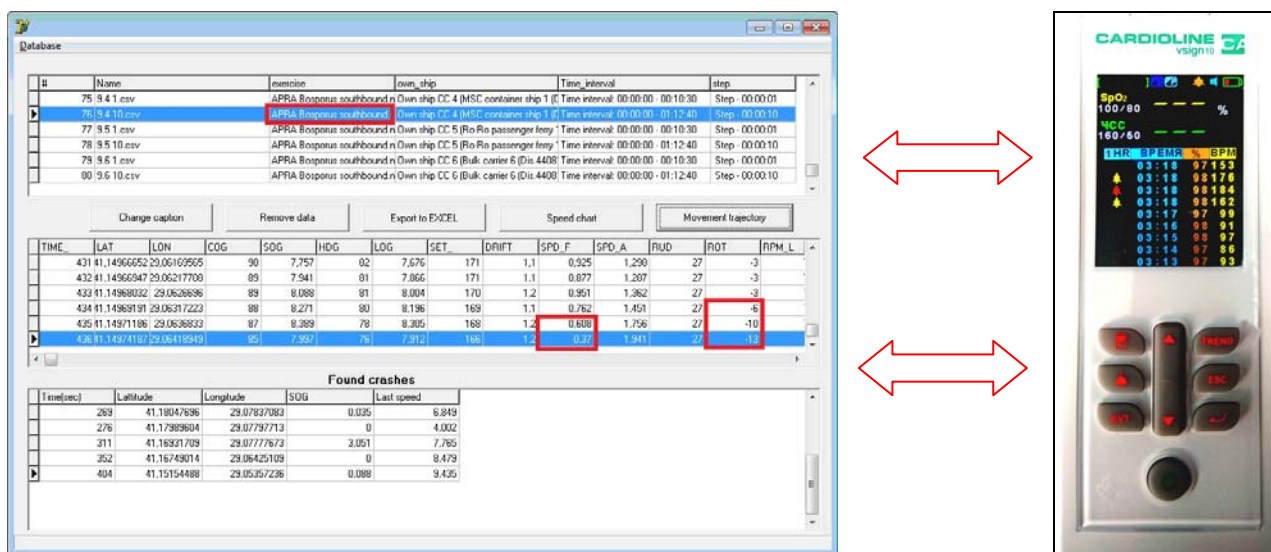


Figure 9 – NTPRO 5000 Simulator Database Analysis

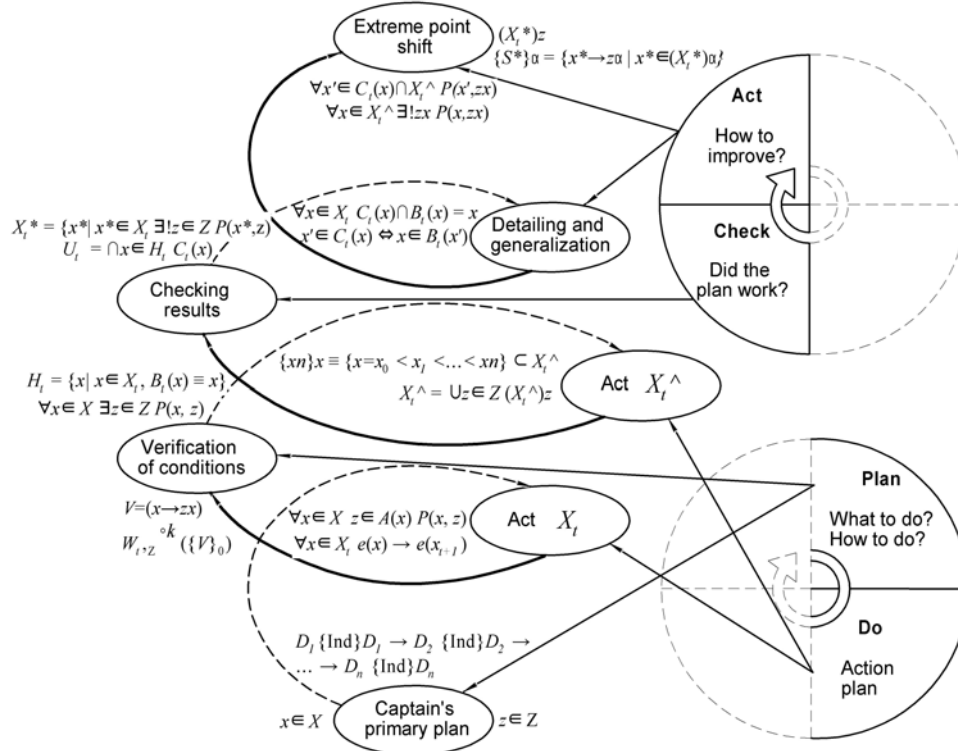


Figure 10 – A model of the dynamics of operators' anticipation centers

Then, the test performance is certain to be depicted by the dependence as follows:

$$Gv(\tau) \mapsto \{Z \text{ Task}\} G, \text{ if } e(\tau/T) = 1 \Rightarrow e(T \text{ Task}) = 1, \\ \text{for } \{Ag\} \mapsto G(\tau) \& \{G(a)\} \mapsto G(\tau).$$

Three-coordinate metric depicts as follows:

- x is the set of previous navigation tasks τ/T ;
- y is the ratio of the diameter of the inscribed triangle with the diameter δ ;
- z is the coefficient of deviation of the triangular surface from the flat on the Cayley graph.

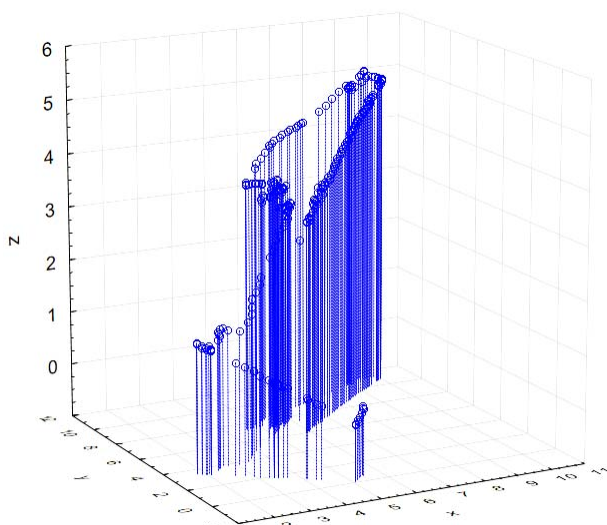


Figure 11 – xyz scatter chart

The scatter diagram (Fig. 11) shows the relationships between the points, with none -speaking of random outliers, having the form of a fragmented trajectory. It confirms the transition hypothesis in accordance with conditions and external factors. The scatter of states mainly witnesses for having all evolutionary increases of the indicators in xyz up.

Let us construct a regression using automated neural nets as it seems to be named as the most effective method for processing the data of the presented dimensions [29]. Suppose, the settings of the samples might be admitted as follows: training is 70%; control and test are both equal to 15% with getting the number of subsamples equal to 12. Let us take a multi-layer perceptron being a type of neural net, 20 nets for training and 5 ones for output as the best issues to be chosen from.

In addition, for comparison purposes we might be using a custom neural network. To do this, the hyperbolic function of network activation is to be chosen singling out 5 networks for training and the number of hidden neurons. Further, for the perceptron the second-order accuracy algorithm BFGS and the number of items is 3000 i reported to be taken (Fig. 12, Fig. 13).

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct
2	MLP 3-6-1	0.866019	0.963812	0.952913	BFGS 137	SOS
3	MLP 3-10-1	0.944001	0.930048	0.964935	BFGS 255	SOS
4	MLP 3-9-1	0.899128	0.973911	0.976353	BFGS 283	SOS

Figure 12 – Best MLP Network 3-9-1

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct.
6	MLP 3-4-1	0.512810	0.280650	0.254278	BFGS_54	SOS
7	MLP 3-4-1	0.895043	0.971929	0.974676	BFGS_29...	SOS
8	MLP 3-4-1	0.883500	0.963011	0.982401	BFGS_139	SOS
9	MLP 3-4-1	0.656596	0.472026	0.573212	BFGS_84	SOS

Figure 13 – Best MLP Network 3-4-1

To obtain the most effective models a neural network basing on the multiple subsampling method is intended to be constructed (Fig. 14). For meeting these requirements the radial basis function and the number of hidden neurons equal to 30 are offered to be taken.

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct.
11	RBF 3-30-1	-0.129662	-0.100111	-0.036909	RBFT	SOS
12	RBF 3-30-1	-0.061876	-0.049586	-0.091447	RBFT	SOS
13	RBF 3-30-1	-0.089920	-0.100461	-0.136163	RBFT	SOS
14	RBF 3-30-1	-0.009433	-0.469965	-0.014456	RBFT	SOS

Figure 14 – Using the radial basis function

As can clearly be seen the radial basis function delivers disproportionately low rates not granting its use for the sample. The total learning outcomes of neural networks are presented on Fig. 15.

Index	Net. name	Training perf.	Test perf.	Validation perf.	Training error	Test error	Validation error
23	MLP 3-4-1	0.885490	0.978961	0.601386	0.000012	0.000006	0.004063
24	MLP 3-4-1	0.595815	0.366749	0.551690	0.000033	0.000094	0.000102
25	MLP 3-4-1	0.647822	0.655474	0.366311	0.000045	0.000062	0.000045
26	MLP 3-4-1	0.422636	0.595017	0.283890	0.000050	0.000018	0.000164
27	MLP 3-4-1	0.935426	0.901746	0.906628	0.000009	0.000019	0.000015
28	MLP 3-4-1	0.832792	0.897238	0.748960	0.000019	0.000050	0.000010
29	MLP 3-4-1	-0.034971	-0.174446	0.389461	0.000052	0.000071	0.000170
30	MLP 3-4-1	0.397231	0.427485	0.019626	0.000075	0.000024	0.000058
31	MLP 3-4-1	0.392804	0.685555	0.017583	0.000071	0.000012	0.000077
32	MLP 3-4-1	0.405597	0.555839	0.497110	0.000077	0.000022	0.000017
33	MLP 3-4-1	0.388823	0.645078	0.351217	0.000068	0.000017	0.000071
34	MLP 3-4-1	0.862818	0.943959	0.449031	0.000019	0.000012	0.000028

Figure 15 – Model Performance Comparison

Suppose, let us conduct a graphical analysis of the adequacy of simulated neural networks by the help of a scatter diagram (Fig. 16), and response variables in the form of a three-dimensional diagram (Fig. 17).

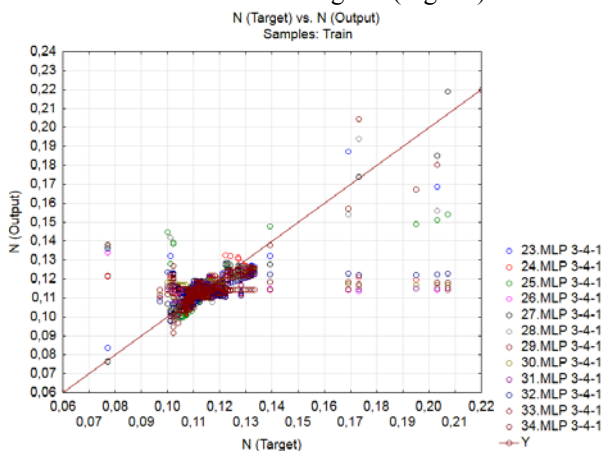


Figure 16 – Scatter Diagram Analysis of the adequacy of simulated neural networks using

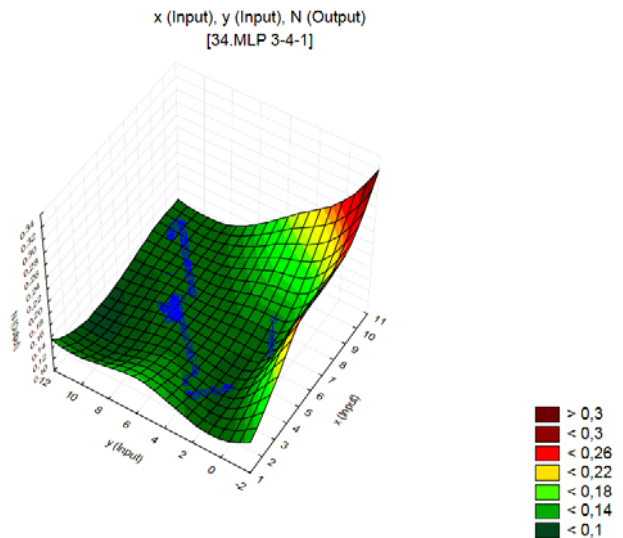


Figure 17 – Three-dimensional diagram of response variables

According to the scatter diagram the accumulation of points is being depicted as a normal one and the number of outliers is being shown a random one. In addition, a three-dimensional reliability diagram is sure to display well enough the dependence of variables and surface showing rather high sufficiency.

In its turn, analysis of the sensitivity of the variables showed that the variable z , “the coefficient of deviation of the triangular surface from the flat on the Cayley graph”, is considered to be the most significant. This point testifies in favor of the hypothesis of the study (Fig. 18).

Networks	Sensitivity analysis (Лист1 in ! Samples: Train		
	z	y	x
23. MLP 3-4-1	20956,2	47,3	27931,83
24. MLP 3-4-1	15,3	19,1	2,18
25. MLP 3-4-1	5605,6	51,8	5,84
26. MLP 3-4-1	7,2	5,9	1,14
27. MLP 3-4-1	294388,4	332621,3	4,13
28. MLP 3-4-1	66736,1	51135,9	492,55
29. MLP 3-4-1	1,0	1,0	1,00
30. MLP 3-4-1	7,7	6,3	1,02
31. MLP 3-4-1	8,2	7,2	1,04
32. MLP 3-4-1	11,9	13,0	1,19
33. MLP 3-4-1	9,3	7,1	1,23
34. MLP 3-4-1	1467,7	258,6	5,29
Average	32434,6	32014,5	2370,70

Figure 18 – The sensitivity of variables

The analysis of the model for predicting conditions turned out to show rather high values. Thus, it becomes possible to create environment likely to escape from having negative manifestations of the human factor of the operator (Fig. 19).

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct.
27	MLP 3-4-1	0.935426	0.901746	0.906628	BFGS_181	SOS

Figure 19 – Prediction of model No. 27 states

Suppose, let us find out the quality of model 27. So, the average value, with the relative error on the test set as being 3.5% (0.0350806479), indicates the high quality of the constructed neural network models (Fig. 20).

Case name	Predictions spreadsheet for N (Лист1 in 1 стат)				
	Network: 27_MLP 3-4-1				
	Samples: Test				
	1				
	Маре				
MEAN case 1-50	0,0350806479				

Figure 20 – Quality analysis of model No. 27

CONCLUSIONS

The conclusion reached is that formal approaches, models and algorithms have succeeded providing a positive result in identification of anticipation of operators in maritime transport.

The results of the experiment using NTPRO 5000, as well as the developed software, have created the environment favorable to identify either the fact of the manifestation of operator anticipation on the decision path nor to track its dynamics over a long period in addition.

Automated identification rests on the hypothesis that the manifestation of anticipation in operators are highly likely to have been implementing during the time of performing complex maneuvers when navigating the marine transport. Reaching the research outcome was possible to be carried out by means of the use of formal approaches and automation within the framework of decision support systems. To meet these results the metric space Y_g with a metric d_{y_g} were constructed to display the manifestations of anticipation of the operator.

In order to form the trajectory of decision-making by the maritime transport operator a three-factor model ERO-AEA-EAPI was proposed to be done on the surface of Cayley graph X_g with a metric d_{x_g} . The synchronization of the data of the NTPRO 5000 navigation simulator enable the identification of the extrema on the Cayley graph of the quasi-isometric surface. The mentioned above identified surface distortion in the form of δ neighborhood surrounding let us find out the operator's anticipation centers.

It must be heighted that the experiment using the Transas navigation simulator NTPRO 5000 have verified the research hypothesis. These issues were profitably coped with thanks to the developed automated module for constructing a triangular space. Besides, identification of the centers, magnitude and dynamics of anticipation of operators were succeeded to be achieved.

Singled out indicators of anticipation x' as well as the influence of new data on the psycho-physiological parameters of the operator are reported to be beneficial contributors to the development of a model of the dynamics of operators' anticipation centers in the form of a spiral Shukhart-Deming cycle. The further tendency in studying was stated as receiving an artificial neural network for the sake of having predictions about the likely negative mani-

festations of anticipation in discrete form by constructing multiple regression.

Thus, this study allowed us to step forward in problem solving of identification and predicting the anticipation of operators in maritime transport by introducing DSS with an aim of creating and enhancing safety environment in maritime transport.

ACKNOWLEDGEMENTS

The work is carried out within the framework of “Development of software for important the quality of functioning systems for dynamic position sea vessels” (state registration number 0119U100948), of navigation and ECDIS departments of Kherson State Maritime Academy Navigation Faculty (scientific adviser: Ph.D. Associate Professor, Deputy Rector for scientific and pedagogical work, Kherson State Maritime Academy, Ukraine, Ben A. P.).

REFERENCES

1. Harati-Mokhtari A., Wall A., Brooks P. at al. Automatic Identification System (AIS): Data Reliability and Human Error Implications, *Journal of Navigation*, 2007, Vol. 3, Issue 60, pp. 373–389. DOI:10.1017/S0373463307004298.
2. Be in Command by Knowing These ECDIS Safety Settings. Written by Capt Rajeev Jassal (on August 20), 2016 [Electronic resource]. Access mode: <https://www.myseatime.com/blog/detail/be-in-command-by-knowing-these-ecdis-safety-settings>.
3. Yurzenko A. An E-course based on the LMS Moodle to teach “Maritime English for Professional purpose”, *Information Technologies and Learning Tools*, 2019, Issue 71, P. 92. DOI: 10.33407/itlt.v71i3.2512.
4. Zinchenko S. M., Ben A. P., Nosov P. S. et al. Improving Accuracy and Reliability in Automatic Ship Motion Control Systems, *Radio Electronics, Computer Science, Control*, 2020, Vol. 2, pp. 183–195.
5. Nosov P. S., Zinchenko S. M., Popovych I. S. et al. Diagnostic system of perception of navigation danger when implementation complicated maneuvers, *Radio Electronics, Computer Science, Control*, 2020, Vol. 1, pp. 146–161. DOI: 10.15588/1607-3274-2020-1-15.
6. Nosov P., Palamarchuk I., Zinchenko S. at al.]Development of means for experimental identification of navigator attention in ergatic systems of maritime transport, *Bulletin of University of Karaganda. Technical Physics*, 2020, Vol. 1, Issue 97, pp. 58–69. DOI: 10.31489/2020Ph1/58-69.
7. Geometric group theory. Maljutin, A. (2020) [Electronic resource]. Access mode: <https://www.lektorium.tv/node/32971>
8. Kononenko O., Kononenko A., Stynska V. et al. Research of the factor structure of the model of world view settings at a young age, *Revista Inclusiones*, 2020, Vol. 7, pp. 98–116.
9. Powell E., Klinck J., Munroe D. et al. The value of captains' behavioral choices in the success of the surfclam (spisula solidissima) fishery on the U.S. Mid-Atlantic Coast: a Model Evaluation, *Journal of Northwest Atlantic Fishery Science*, 2020, Vol. 47, pp. 1–27. DOI: 10.2960/J.v47.m701.
10. Pinheiro L., Fortes C., Santos J. Risk Analysis and Management of Moored Ships in Ports, *37th International Conference on Ocean, Offshore and Arctic Engineering, Madrid, 17–22 June 2018: proceedings*. ASME, 2018. V003T02A057. DOI:10.1115/OMAE2018-78396

11. Chen L., Rong-chang C. Current problems and risk analysis on high pollution risk operation of ship, *5th International Conference on Transportation Information and Safety (ICTIS)*. Liverpool, 2019, pp. 136–140. DOI:10.1109/ICTIS.2019.8883804.
12. Gluver H., Olsen D. Current practice in risk analysis of ship collisions to bridges, *Ship Collision Analysis: book*. Copenhagen, 2018, pp. 85–96 DOI:10.1201/9780203739778-9.
13. Lee H-W., Rhyu H-I. Antiaging strategy considering physiological characteristics, *Journal of Exercise Rehabilitation*. – 2019, Vol. 15, pp. 346–350. DOI:10.12965/jer.1938214.107.
14. Yu H., Guo D. Study on Physiological Characteristics of Emotion, *Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*. Qinhuaangdao, 2015, pp. 1286–1289, DOI:10.1109/IMCCC.2015.276.
15. Kurdel P., Lazar T., Labun J. Computing skills in aviation ergatic systems, *12th International Symposium on Applied Machine Intelligence and Informatics (SAMII): IEEE*. Herl'any, 2014, pp. 375–379. DOI: 10.1109/SAMI.2014.6822443.
16. Millán P., Inglada-Pérez L., Casares P. et al. Modelización del transporte marítimo de contenedores, *Estudios de Economía Aplicada*, 2019, Vol. 36, pp. 675. DOI: 10.25115/eea.v36i3.2548.
17. Ertogan M., Tayyar G., Wilson P. et al. Marine measurement and real-time control systems with case studies, *Ocean Engineering*, 2018, Vol. 159, P. 457–469. DOI: 10.1016/j.oceaneng.2018.02.001.
18. Wang X., Fu T., Zhang Y. et al. Information Interface of Artificial Intelligence Medical Device Information, *Big Data Analytics for Cyber-Physical System in Smart City*, 2020, Vol. 1117, pp. 1293–1303. DOI: 10.1007/978-981-15-2568-1_180.
19. Zinchenko S. M., Nosov P. S., Mateychuk V. M. et al. Automatic collision avoidance with multiple targets, including maneuvering ones, *Radio Electronics, Computer Science, Control*, 2019, Vol. 4, pp. 211–221. DOI: 10.15588/1607-3274-2019-4-20.
20. Zinchenko S., Nosov P., Mateichuk V. et al. Automatic collision avoidance system with many targets, including maneuvering ones, *Bulletin of University of Karaganda. Technical Physics*, 2019, Vol. 4, pp. 69–79. DOI: 10.31489/2019Ph4/69-79.
21. Patané I., Cardinali L., Salemme R. et al.]Action Planning Modulates Peripersonal Space, *Journal of Cognitive Neuroscience*, 2018, Vol. 31, pp. 1–14. DOI: 10.1162/jocn_a_01349.
22. Vişan M., Ionita A., Filip F. Data Analysis in Setting Action Plans of Telecom Operators, *Challenges and Applications. Studies in Computational Intelligence*. Springer, 2020, Vol. 869. DOI: 10.1007/978-3-030-39250-5_6.
23. Malyutin A. Pretrees and Arborescent Convexities, *Journal of Mathematical Sciences*, 2016, Vol. 212, pp. 566–576. DOI: 10.1007/s10958-016-2689-1.
24. Gromov M., Gersten S. M. ed. Hyperbolic groups. Essays in group theory, *MSRI Publ.* 8. Springer-Verlag, 1987, pp. 75–263.
25. Rips E., Sela Z. Structure and rigidity in hyperbolic groups, *Geom. Funct. Anal.*, 1994, No. 3, pp. 337–371.
26. Prokopchuk Y.A. Sketch of the Formal Theory of Creativity. Dnepr, PSACEA Press, 2017, 452 p.
27. Lerche J., Neve H., Wandahl S. et al. Combining Takt and Deming Cycles at Operator Level – Practical Study, *The 10th International Conference on Engineering, Project, and Production Management, January*, 2020, pp. 623–632.
28. Lymariiev I. O., Subbotin S. A., Oliinyk A. A. et al. Method of large-scale signals transformation for diagnosis in neural network models, *Radio Electronics, Computer Science, Control*, 2018, Vol. 4, pp. 121–134. DOI: 10.15588/1607-3274-2018-4-11
29. Subbotin S. A. Methods of synthesis of models of quantitative dependencies in the basis of trees of regression, realizing cluster – regression approximation by precedents, *Radio Electronics, Computer Science, Control*, 2019, Vol. 3, pp. 76–85. DOI: 10.15588/1607-3274-2019-3-9

Received 26.06.2020.
Accepted 22.08.2020.

УДК 004.942: 316.454.54

АВТОМАТИЗОВАНА ІДЕНТИФІКАЦІЯ АНТИЦИПАЦІЇ ОПЕРАТОРА НА МОРСЬКОМУ ТРАНСПОРТІ

Носов П. С. – канд. техн. наук, доцент кафедри судноводіння та електронних навігаційних систем Херсонської державної морської академії, Україна.

Попович І. С. – д-р наук, професор кафедри загальної та соціальної психології Херсонського державного університету, Україна.

Чернявський В. В. – д-р наук, професор, ректор Херсонської державної морської академії, Україна.

Зінченко С. М. – канд. техн. наук, старший викладач кафедри управління судном, завідувач лабораторією електронних симуляторів Херсонської державної морської академії, Україна.

Прокопчук Ю. О. – д-р техн. наук, доцент, провідний науковий співробітник відділу системного аналізу та проблем управління, Інститут технічної механіки, Національна академія наук, Україна.

Макарчук Д. В. – канд. техн. наук, старший викладач кафедри управління судном Херсонської державної морської академії, Україна.

АНОТАЦІЯ

Актуальність. У статті розглядається підхід ідентифікації антиципації як важливої складової процесу прийняття рішень оператора на прикладі капітана далекого плавання в ергатичних системах критичних інфраструктур при управлінні морським транспортом. Розглянутий аспект антиципації операторів є складною формою людино-машинного взаємодії і вимагає розробки підходів та відповідних інформаційних засобів.

Мета. Підхід щодо розробки засобів ґрунтується на інформаційному аналізі повного спектра траєкторій прийняття рішень операторів в моменти виконання складних багатоступінних дій заснованих на прийнятій стратегії людино-машинної взаємодії. Такий підхід передбачає формування метрики здатної алгоритмічним способом відобразити простір варіантів в умовах комбінаторного передувалення в термінах геометричної теорії груп на графі Келі.

Метод. В рамках розробки підходу визначена математична модель збору та обробки експериментальних-тальних даних в ході аналізу бази даних навігаційних симуляторів. З метою підтвердження прийнятих формально-алгоритмічних підходів проведено імітаційне моделювання що формує траєкторію прийняття рішень оператора в критичних ситуаціях заснованих © Nosov P. S., Popovych I. S., Cherniavskiy V. V., Zinchenko S. M., Prokopchuk Y. A., Makarchuk D. V., 2020
DOI 10.15588/1607-3274-2020-3-15

на трехфакторной моделі ERO-AEA-EAPI. Таким чином алгебраїчне і програмне представлення метричного простору прийняття рішень дозволяє апроксимувати складні людино-машинні взаємодії в умовах невизначеності і отримати підхід трансформації даних в знання про суб'єкта критичної інфраструктури – оператора.

Результати. З метою підтвердження доцільності розроблених формально-алгоритмічних підходів був проведений експеримент з використанням навігаційного симулятора Navi Trainer 5000 (NTPRO 5000). В ході експерименту були проаналізовані дані сервера при виконанні однієї з найбільш складних операцій – швартуванні, і сформовані знання про антиципацію операторів у вигляді триангулярних конструктів на квазіізотричному просторі графа Келі. Результати моделювання із застосуванням автоматизованої нейронної мережі дозволили отримати множинну регресію даних і провести аналіз зв'язків незалежних змінних про що свідчать результати діаграм розсіювання і надійності.

Висновки. Формально-алгоритмічні підходи, представлені в дослідженні, в поєднанні з розробленими програмними засобами, підходи перетворення даних у метаданні про антиципацію операторів дозволили класифікувати і визначити індивідуальні стратегії прийняття рішень при управлінні судном що дало змогу прогнозувати ймовірність катастрофічних наслідків. Результативність запропонованого підходу і моделей була успішно обґрунтована за результатами автоматизованої обробки експериментальних даних.

КЛЮЧОВІ СЛОВА: системи підтримки прийняття рішень; системи ідентифікації знань; модель оператора; комп'ютерні навігаційні симулятори; управління ризиками; аналіз людського фактора; автоматизовані системи управління.

УДК 004.942: 316.454.54

АВТОМАТИЗИРОВАННАЯ ИДЕНТИФИКАЦИЯ АНТИЦИПАЦИИ ОПЕРАТОРА НА МОРСКОМ ТРАНСПОРТЕ (RU)

Носов П. С. – канд. техн. наук, доцент кафедры судовождения и электронных навигационных систем Херсонской государственной морской академии, Украина.

Попович И. С. – д-р наук, профессор кафедры общей и социальной психологии Херсонского государственного университета, Украина.

Чернявский В. В. – д-р наук, профессор, ректор Херсонской государственной морской академии, Украина.

Зинченко С. М. – канд. техн. наук, старший преподаватель кафедры управления судном, заведующий лабораторией электронных симуляторов Херсонской государственной морской академии, Украина.

Прокопчук Ю. А. – д-р техн. наук, доцент, ведущий научный сотрудник отдела системного анализа и проблем управления, Институт технической механики, Национальная академия наук, Украина.

Макарчук Д. В. – канд. техн. наук, дальнего плавания, старший преподаватель кафедры управления судном Херсонской государственной морской академии, Украина.

АННОТАЦИЯ

Актуальность. В статье рассматриваются подходы идентификации антиципации как важной составляющей процесса принятия решений оператором на примере капитана дальнего плавания в эргатических системах критических инфраструктур при управлении морским транспортом. Рассматриваемый аспект антиципации операторов является сложной формой человеко-машинного взаимодействия и требует разработки информационных подходов и средств.

Цель. Подход к разработке подходов основывается на информационном анализе полного спектра траекторий принятия решений операторов в моменты выполнения сложных многоэтапных действий основанных на принятой стратегии человеко-машинного взаимодействия. Такой подход предполагает формирование метрики способной алгоритмическим способом отобразить пространство вариантов, в условиях комбинаторного представления, в терминах геометрической теории групп на графе Келли.

Метод. В рамках разработки подхода определена математическая модель сбора и обработки экспериментальных данных в ходе анализа базы данных навигационных симуляторов. С целью подтверждения принятых формально-алгоритмических подходов проведено имитационное моделирование формирующее траекторию принятия решений оператора в критических ситуациях основанных на трехфакторной модели ERO-AEA-EAPI. Таким образом алгебраическое и программное представление метрического пространства принятия решений позволяет аппроксимировать сложные человеко-машинные взаимодействия в условиях неопределенности и получить подход трансформации данных в знания о субъекте критической инфраструктуры – операторе.

Результаты. С целью подтверждения целесообразности разработанного подхода и формально-алгоритмических подходов был проведен эксперимент с использованием навигационного симулятора Navi Trainer 5000 (NTPRO 5000). В ходе эксперимента были проанализированы данные сервера в ходе выполнения одной из наиболее сложных операций – швартовки, и сформированные знания об антиципации в виде триангулярных конструктов на квазиизотрическом пространстве графа Келли. Результаты моделирования с применением автоматизированных нейронных сетей позволили получить множественную регрессию данных и провести анализ связей множества независимых переменных о чем свидетельствуют результаты диаграмм рассеивания и надежности.

Выводы. Формально-алгоритмические подходы, представленные в исследовании, в сочетании с разработанными программными средствами, подходы преобразования данных в знания об антиципации операторов позволили классифицировать и определить индивидуальные стратегии принятия решений при управлении судном и прогнозировать вероятность катастрофических последствий. Результативность предложенного подхода и моделей была успешно обоснована по результатам автоматизированной обработки экспериментальных данных.

КЛЮЧЕВЫЕ СЛОВА: системы поддержки принятия решений; системы идентификации знаний; модель оператора; компьютерные навигационные симуляторы; управление рисками; анализ человеческого фактора; автоматизированные системы управления.

ЛІТЕРАТУРА / LITERATURA

1. Automatic Identification System (AIS): Data Reliability and Human Error Implications / [A. Harati-Mokhtari, A. Wall, P. Brooks et al.] // *Journal of Navigation*. – 2007. – Vol. 3, Issue 60. – P. 373–389. DOI:10.1017/S0373463307004298.
2. Be in Command by Knowing These ECDIS Safety Settings. Written by Capt Rajeev Jassal (on August 20), 2016 [Electronic resource]. – Access mode: <https://www.myseatime.com/blog/detail/be-in-command-by-knowing-these-ecdis-safety-settings>.
3. Yurzhenko A. (2019). An E-course based on the LMS Moodle to teach “Maritime English for Professional purpose” / A. Yurzhenko // *Information Technologies and Learning Tools*. – 2019. – Issue 71. – P. 92. DOI: 10.33407/itlt.v71i3.2512.
4. Improving Accuracy and Reliability in Automatic Ship Motion Control Systems / [S. M. Zinchenko, A. P. Ben, P. S. Nosov et al.] // *Radio Electronics, Computer Science, Control*. – 2020. – Vol. 2. – P. 183–195.
5. Diagnostic system of perception of navigation danger when implementation complicated maneuvers / [P. S. Nosov, S. M. Zinchenko, I. S. Popovych et al.] // *Radio Electronics, Computer Science, Control*. – 2020. – Vol. 1. – P. 146–161. DOI: 10.15588/1607-3274-2020-1-15.
6. Development of means for experimental identification of navigator attention in ergatic systems of maritime transport / [P. Nosov, I. Palamarchuk, S. Zinchenko et al.] // *Bulletin of University of Karaganda. Technical Physics*. – 2020. – Vol. 1, Issue 97. – P. 58–69. DOI: 10.31489/2020Ph1/58-69.
7. Geometric group theory. Malyutin, A. (2020) [Electronic resource]. – Access mode: <https://www.lektorium.tv/node/32971>
8. Research of the factor structure of the model of world view settings at a young age / [O. Kononenko, A. Kononenko, V. Stynska et al.] // *Revista Inclusiones*. – 2020. – Vol. 7. – P. 98–116.
9. The value of captains’ behavioral choices in the success of the surfclam (*spisula solidissima*) fishery on the U.S. Mid-Atlantic Coast: a Model Evaluation / [E. Powell, J. Klinck, D. Munroe et al.] // *Journal of Northwest Atlantic Fishery Science*. – 2020. – Vol. 47. – P. 1–27. DOI: 10.2960/J.v47.m701.
10. Pinheiro L. Risk Analysis and Management of Moored Ships in Ports / L. Pinheiro, C. Fortes, J. Santos // *37th International Conference on Ocean, Offshore and Arctic Engineering, Madrid, 17–22 June 2018: proceedings*. – ASME, 2018. V003T02A057. DOI:10.1115/OMAE2018-78396
11. Chen L. Current problems and risk analysis on high pollution risk operation of ship / L. Chen, C. Rong-chang // *5th International Conference on Transportation Information and Safety (ICTIS)*. – Liverpool. – 2019. – P. 136–140 DOI:10.1109/ICTIS.2019.8883804.
12. Gluver H. Current practice in risk analysis of ship collisions to bridges / H. Gluver, D. Olsen // *Ship Collision Analysis: book*. – Copenhagen, 2018. – P. 85–96 DOI:10.1201/9780203739778-9.
13. Lee H-W. Antiaging strategy considering physiological characteristics / H-W. Lee, H-I. Rhyu // *Journal of Exercise Rehabilitation*. – 2019. – Vol. 15. – P. 346–350. DOI:10.12965/jer.1938214.107.
14. Yu H. Study on Physiological Characteristics of Emotion / H. Yu, D. Guo // *Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*. – Qinhuangdao. – 2015. – P. 1286–1289. DOI:10.1109/IMCCC.2015.276.
15. Kurdel P. Computing skills in aviation ergatic systems / P. Kurdel, T. Lazar, J. Labun // *12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*: IEEE. – Herl’any. – 2014. – P. 375–379. DOI: 10.1109/SAMI.2014.6822443.
16. Modelización del transporte marítimo de contenedores / [P. Millán, L. Inglada-Pérez, P. Casares et al.] // *Estudios de Economía Aplicada*. – 2019. – Vol. 36. – P. 675. DOI: 10.25115/eea.v36i3.2548.
17. Marine measurement and real-time control systems with case studies / [M. Ertogan, G. Tayyar, P. Wilson et al.] // *Ocean Engineering*. – 2018. – Vol. 159. – P. 457–469. DOI: 10.1016/j.oceaneng.2018.02.001.
18. Information Interface of Artificial Intelligence Medical Device Information / [X. Wang, T. Fu, Y. Zhang et al.] // *Big Data Analytics for Cyber-Physical System in Smart City*. – 2020. – Vol. 1117. – P. 1293–1303. DOI: 10.1007/978-981-15-2568-1_180.
19. Automatic collision avoidance with multiple targets, including maneuvering ones / [S. M. Zinchenko, P. S. Nosov, V. M. Mateychuk et al.] // *Radio Electronics, Computer Science, Control*. – 2019. – Vol. 4. – P. 211–221. DOI: 10.15588/1607-3274-2019-4-20.
20. Automatic collision avoidance system with many targets, including maneuvering ones / [S. Zinchenko, P. Nosov, V. Mateichuk et al.] // *Bulletin of University of Karaganda. Technical Physics*. – 2019. – Vol. 4. – P. 69–79. DOI: 10.31489/2019Ph4/69-79.
21. Action Planning Modulates Peripersonal Space / [I. Patané, L. Cardinali, R. Salemme et al.] // *Journal of Cognitive Neuroscience*. – 2018. – Vol. 31. – P. 1–14. DOI: 10.1162/jocn_a_01349.
22. Vişan, M. Data Analysis in Setting Action Plans of Telecom Operators / M. Vişan, A. Ionita, F. Filip // *Challenges and Applications. Studies in Computational Intelligence*. Springer. – 2020. – Vol. 869. DOI: 10.1007/978-3-030-39250-5_6.
23. Malyutin A. Pretrees and Arborescent Convexities / A. Malyutin // *Journal of Mathematical Sciences*. – 2016. – Vol. 212. – P. 566–576. DOI: 10.1007/s10958-016-2689-1.
24. Gromov M. Hyperbolic groups. *Essays in group theory* / Gromov M. (S. M. Gersten, ed.) // *MSRI Publ.* 8. – Springer-Verlag, 1987. – P. 75–263.
25. Rips E. Structure and rigidity in hyperbolic groups / E. Rips, Z. Sela. – *Geom. Funct. Anal.* – 1994. – No. 3. – P. 337–371.
26. Prokopchuk Y.A. Sketch of the Formal Theory of Creativity / Y. A. Prokopchuk. – Dnepr : PSACEA Press, 2017. – 452 p.
27. Combining Takt and Deming Cycles at Operator Level – Practical Study / [J. Lerche, H. Neve, S. Wandahl et al.] // *The 10th International Conference on Engineering, Project, and Production Management, January 2020*. – P. 623–632.
28. Method of large-scale signals transformation for diagnosis in neural network models / [I. O. Lymariev, S. A. Subbotin, A. A. Oliinyk et al.] // *Radio Electronics, Computer Science, Control*. – 2018. – Vol. 4. – P. 121–134. DOI: 10.15588/1607-3274-2018-4-11
29. Subbotin S. A. Methods of synthesis of models of quantitative dependencies in the basis of trees of regression, realizing cluster – regression approximation by precedents / S. A. Subbotin // *Radio Electronics, Computer Science, Control*. – 2019. – Vol. 3. – P. 76–85. DOI: 10.15588/1607-3274-2019-3-9

ПРОГРАМНИЙ АНАЛІЗ СТРУКТУРИ БЛОЧНО-ЦИКЛІЧНОЇ БАЗИСНОЇ МАТРИЦІ ДКП

Процько І. О. – д-р техн. наук, доцент кафедри автоматизованих систем управління Національного університету «Львівська політехніка», Львів, Україна.

Міщук М. В. – студент кафедри автоматизованих систем управління Національного університету «Львівська політехніка», Львів, Україна.

АНОТАЦІЯ

Актуальність. Використання матричного запису застосовується для формалізації предметної галузі в рамках алгебраїчного підходу. Ефективне обчислення дискретних косинусних перетворень використовує приведення гармонічного базису до блочно-циклічних матричних структур з подальшим обчисленням перетворень за допомогою швидких циклічних згорток. Детальний аналіз структури базисної блочної матриці перетворення забезпечує синтез ефективних алгоритмів дискретних косинусних перетворень довільних обсягів. Програмна реалізація проведення аналізу блочно-циклічних структур формує опис структури, що дозволяє зменшити обчислювальну складність алгоритму та виконати розпаралелення обчислень циклічних згорток.

Мета роботи – визначення алгоритмічних особливостей проведення аналізу структури базисної блочно-циклічної матриці, яка містить цілочисельні аргументи базисних гармонічних функцій, що дозволить зменшити обчислювальну складність синтезованого алгоритму дискретного косинусного перетворення на основі циклічних згорток.

Метод. Пошук і аналіз перебором елементів матриці зі змінним кроком, що враховує властивості блочності та циклічності сформованої базисної матриці дискретного косинусного перетворення, дозволяє швидше виконувати аналіз структури базисної блочної матриці перетворення в порівнянні з повним скануванням.

Результати. Розроблено алгоритмічне та програмне забезпечення виконання аналізу структури блочно-циклічної базисної матриці, за допомогою якого визначається масив параметрів даних формального опису структури базису дискретного косинусного перетворення. Дані аналізу структури базисної матриці дозволяють визначити наявність ідентичних циклічних підматриць розміщених горизонтально або вертикально одна відносно іншої і, тим самим, зменшити кількість виконання циклічних згорток.

Висновки. Ефективний аналіз блочно-циклічної структури базисної матриці на основі розробленого програмного забезпечення є важливою частиною процесу синтезу швидкого алгоритму, який забезпечує скорочення обчислювальної складності та можливість розпаралелення виконання дискретного косинусного перетворення. Розроблене алгоритмічне та програмне забезпечення виконання аналізу структури блочно-циклічної базисної матриці може застосовуватись і для аналізу структури та пошуку відповідних підматриць в будь-яких матрицях з цілими, дійсними та нульовими елементами.

КЛЮЧОВІ СЛОВА: аналіз матриці, алгоритм пошуку, твірний масив, блочно-циклічні підматриці, дискретні косинусні перетворення.

АБРЕВІАТУРИ

БЦС – блочно-циклічні структури;
ДКП – дискретне косинусне перетворення;
ТМ – твірний масив;
ЦЗ – циклічні згортки.

W_N – базисна матриця перетворення;
 $\gamma_{k,n}$ – відповідні масштабувальні множники.

ВСТУП

Запис даних у вигляді k -мірної матриці $A(n_1, n_2, \dots, n_k)$, де кожен вимір може відповідати k параметрам пов'язаним зі значенням матричного елемента, використовується для формалізації предметної галузі в рамках алгебраїчного підходу. Використання матричного запису застосовується не тільки в системах лінійних рівнянь та лінійних перетвореннях, але і в різноманітних прикладних задачах інтелектуального аналізу даних, таких як дослідження генетичних послідовностей, текстової і веб-інформації, риночних кошків, розпізнавання образів та інших. В процесі розробки ефективних алгоритмів розв'язку цих задач виконується приведення матриці її до відповідного вигляду за допомогою факторизації або декомпозиції та подальшого аналізу структури сформованої матриці.

Термін аналіз структури матриці достатньо багатогранний, який визначається як алгебраїчним підходом, так і формальним специфічним математичним апаратом конкретної предметної області. За формою

НОМЕНКЛАТУРА

a – зміщення;
 $A(n_1, n_2, \dots, n_k)$ – k -мірна матриця;
 b – зміщення;
 h_{ij} – елемент твірного під масиву;
 $H(n)$ – твірний масив;
 $H_i(L_i)$ – твірний підмасив;
 k – число підмасивів у твірному масиві;
 L – обсяг твірного масиву;
 L_i – обсяг під масиву;
 N – обсяг перетворення;
 T – період базисної функції;
 $x(n)$, вхідна послідовність перетворення;
 $X^c(m)$ – вихідна послідовність перетворення;
 v_{ij} – елементи матриці аргументів;
 \bar{v}_{ij} – спрощені елементи матриці аргументів;
 $V_N(n_1, n_2)$ – матриця спрощених аргументів;

виділяють діагональні і трикутні матриці, а за основними властивостями рядків і стовпців бувають симетричні або кососиметричні, розріджені матриці, латинські та магічні квадрати та інші. Актуальним для організації ефективних обчислень є застосування блочних матриць [1]. Розділення матриці на блоки полягає у повному її поділі на неперетинаючі підматриці (блоки), де будь-який її елемент належить одній і тільки одній з цих підматриць. Наприклад, кронекерівський добуток матриць в результаті формує блочну матрицю. В блочній матриці підматриці можуть бути діагональними, трикутними, а сама блочна матриця може мати блочно-діагональну структуру, циклічно-блочну та інше [2].

Одним з підходів ефективного обчислення ДКП є приведення гармонічного базису до блочно-циклічних матричних структур з подальшим обчисленням перетворень за допомогою швидких циклічних згорток [3]. В роботі [4] для приведення гармонічного базису перетворення ДКП до набору циклічних зліва підматриць використовуються твірні масиви. Подальший синтез швидкого алгоритму перетворення потребує аналізу одержаної БЦС з метою скорочення обчислювальної складності та ефективної організації його виконання.

Об'єктом дослідження є алгоритмічний та програмний процес виконання аналізу структури цілочисельної блочної матриці, яка містить аргументи базисних гармонічних функцій ДКП.

Предметом дослідження є особливості визначення параметрів, які характеризують БЦС базисної матриці ДКП.

Метою роботи є визначення алгоритмічних особливостей виконання аналізу структури блочно-циклічної базисної матриці, яка містить цілочисельні аргументи базисних гармонічних функцій ДКП.

1 ПОСТАНОВКА ЗАДАЧІ

Підхід ефективного обчислення різних видів ДКП I–IV, що базується на декомпозиції матриці аргументів базисної функції розглянуто в роботах [4, 5]. В результаті застосування підходу, структуру базисної матриці можна задати ТМ

$$H(L) = H_1(L_1)H_2(L_2)\dots H_k(L_k) = (h_{11}, h_{12}, \dots, h_{1L_1})(h_{21}, h_{22}, \dots, h_{2L_2})\dots (h_{kL_1}, h_{kL_2}, \dots, h_{kL_k}). \quad (1)$$

Обсяг твірною масиву L дорівнює сумі обсягів підмасивів L_i

$$L = (L_1 + L_2 + \dots + L_k). \quad (2)$$

де h_{ij} – цілочисельні елементи твірною підмасиву $H_i(L_i)$, які є аргументами базисної гармонічної функції перетворення ($i=1,2,\dots,k; j=1,2,\dots,L_i$) та $h_{ij} < T/2$ менші періоду повторення базисної гармонічної функції; а k – число підмасивів у твірною масиві $H(n)$, яке визначається конкретним значенням N обсягу перетворення. Для основних видів ДКП I–IV, що знаходять широке застосування в процесі обробки інформацій-

них даних та сигналів, подано в табл. 1 значення періодичності та обсягу ТМ.

Таблиця 1– Значення періодичності та обсягу ТМ основних видів ДКП

Вид перетворення	ДКП–I	ДКП–II	ДКП–III	ДКП–IV
Періодичність, T	$2N$	$4N$	$4N$	$8N$
Обсяг ТМ, L	N	$2N$	$2N$	$2N$

ТМ визначається за підстановкою з рядків/стовпців аргументів гармонічної функції базисної матриці W_N перетворення обсягу N [6]. На основі твірною масиву здійснюється переіндексація рядків/стовпців базисної матриці, що в результаті приводить до формування БЦС в базисній матриці для різних видів ДКП.

Отже, одержуємо матрицю $V_N(n_1, n_2)$, що містить набір цілочисельних циклічних зліва підматриць різних обсягів, де розмірності n_1, n_2 визначаються обсягом твірною масивів для рядків та стовпців. Кожна з циклічних квадратних підматриць містить цілочисельні елементи, які належать одному з твірною підмасивів $H_i(L_i)$. Обсяги L_i кожного з твірною підмасивів є довільними, що залежать від обсягу перетворення N , і в сумі відповідають умові (2).

Для синтезу швидкого алгоритму перетворення необхідно виконати аналіз структури одержаної блочно-циклічної матриці з метою визначення ідентичних блоків, що розміщених горизонтально та вертикально один відносно іншого. Наявність ідентичних блоків приводить до зменшення обчислювальної складності та надає можливість розпаралелення виконання обчислення ДКП.

2 ОГЛЯД ЛІТЕРАТУРИ

Аналіз структури матриці широко використовують в процесі аналізу зображень. В більшості графічних застосунків, в якості значень двовимірної матриці, використовується освітленість Y , що дорівнює зваженій сумі $Y=0,3R+0,59G+0,11B$. Ваги вибирають так, щоб максимально точно відповідати чутливості ока до R -червоного, G -зеленого і B -синього. Аналіз елементів освітленості в матриці і відповідно її структури залежить від схем оцінки пікселя в конкретному застосунку [7].

Актуальними є дослідження аналізу структури розріджених матриць. Для цього використовуються високопродуктивні обчислювальні апаратні [8] та програмні [9] засоби, нейронні мережі [10], що дозволяють класифікувати розріджені матриці, виділяти групи ненульових елементів, які можуть оброблятися незалежно. За результатами аналізу будується декомпозиція даних, що забезпечить найкоротший час розрахунків для конкретної структури матриці.

Таким чином, аналіз структури матриці залежить від застосування, яке накладає ряд вимог та критеріїв проведення аналізу. Для синтезу ефективних алгоритмів обчислення ДКП на основі циклічних згорток аналіз базисної матриці зводиться до пошуку ідентич-

них циклічних зліва цілочисельних підматриць. Критерієм та вихідними параметрами аналізу блочно-циклічних матриць є визначення ідентичних блоків, відповідних індексів розташування та належність елементів до одного з твірних підмасивів.

3 МАТЕРІАЛИ І МЕТОДИ

В роботах [4, 11] описано методику та особливості обчислення твірних масивів (1) для синтезу швидких алгоритмів ДКП I–IV.

В загальному перетворення ДКП I–IV задається формулою:

$$X^c(m) = \gamma_{m,n} \sum_{n=0}^{N-1} x(n) \cos \left[\frac{(2m+a)(2n+b)\pi}{4N} \right], \quad (3)$$

де $\gamma_{m,n}$ – відповідні масштабувальні множники, a, b – зміщення в загальному виразі (3) можуть набувати значення $a, b \in \{0, 1\}$ і визначають види перетворень, $x(n)$, $X^c(m)$ – вхідна та вихідна послідовності перетворення ($m, n=0, 1, \dots, N-1$).

На основі елементів твірною масиву $H(L)$ здійснюється переіндексація рядків/стовпців базисної матриці, що в результаті приводить до формування БЦС в базисній матриці.

Наприклад, у ДКП-I, для якого $a=b=0$ в загальному виразі (3), обсягу $N=15$ за твірним масивом $H(14)=(1,2,4,8)(14,13,11,7)(3,6,12, 9)(5,10)$, відповідно методиці, БЦС матриці $V_{15}(14, 14)$ матиме вигляд

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

В роботі [6] описано алгоритм визначення твірною масиву. Особливістю цього комбінаційного алгоритму є те, що для формування циклічного розкладу підстановки використовується відбір перших елементів циклічних підмасивів в зростаючому порядку. Наприклад, для ДКП-II, для якого $a=0$ і $b=1$ в загальному виразі (3), обсягу $N=14$ ТМ має вигляд

$$H(28)=(0)(1,3,9,27,25,19)(2,6,18)(4,12,20)(5,15,11,23,13,17)(7,21)(8,24,16)(10,26,22)(14).$$

Однак, для відбору ідентичних блоків є важливим для аналізу наявність однакових перших елементів циклічних підмасивів і, відповідно, послідовності рядків. У кожному підмасиві твірною масиву, що сфор-

мовано в результаті циклічного розкладу підстановки, є можливим циклічний зсув елементів. Тому для забезпечення чіткого визначення ідентичних підматриць, розміщених горизонтально та вертикально одна відносно інших, та наглядної структуризації блочно-циклічної матриці необхідно перерозподілити твірні підмасиви $(h_{i1}, h_{i2}, \dots, h_{iLi})$ твірною масиву

$$H(L) = (h_{i1}, h_{i2}, \dots, h_{iLi}) (h_{(i+1)1}, \dots, h_{(i+1)L(i+1)}), \quad (4)$$

$$(i = 1, 2, \dots, k; j = 1, 2, \dots, Li),$$

в якому $h_{(i+1)1}$ – перший цілочисельний елемент наступного твірною підмасиву $H_{i+1}(L_{i+1})$ визначається відповідно умови

$$h_{(i+1)1} = \begin{cases} L - h_{i1}, & \text{якщо } h_{i1} < L; \\ h_{(i+1)1}, & \text{якщо } h_{i1} < h_{(i+1)1} < h_{(i+2)1}. \end{cases} \quad (5)$$

Ця умова враховує те, що відповідно методиці [4], базисна матриця містить множину спрощених аргументів базисної функції ДКП. Значення спрощених елементів \bar{v}_{ij} в матриці аргументів ДКП I–IV, відповідно властивостям періодичності та симетричності базисних функцій, визначаються послідовністю обчислень:

$$v_{ij} = T - [(v_{ij}) \bmod T], \text{ якщо } [(v_{ij}) \bmod T] > T/2; \quad (6)$$

$$\bar{v}_{ij} = T/2 - (T - v_{ij}), \text{ якщо } (T - v_{ij}) > T/4 \text{ або } v_{ij} > T/4; \quad (7)$$

інакше $\bar{v}_{ij} = v_{ij}$, ($i = 1, 2, \dots, n_1; j = 1, 2, \dots, n_2$).

Приклад перерозподілу твірних підмасивів у твірною масиві для ДКП-II обсягу $N=14$, який в результаті виконання умови (5) має вигляд

$$H(28)=(0)(1,3,9,27,25,19)(2,6,18)(26,22,10)(4,12,20)(24,16,8)(5,15,11,23,13,17)(7,21)(14).$$

На основі отриманого твірною масиву здійснюється переіндексація рядків/стовпців аргументів функції базисної матриці, що в результаті приводить до формування БЦС в базисній матриці.

Для дослідження базисних матриць ДКП розроблено універсальний програмний засіб аналізу цілочисельних матриць, який виконує сканування всього набору елементів матриці. Для пошуку заданого фрагменту за максимальною шириною і висотою, відповідно його розмірності, в матриці виконується покрокове сканування з переміщенням згори вниз і зліва направо. Результат аналізу цілочисельних матриць і виділення еталонних фрагментів подано на рис. 1.

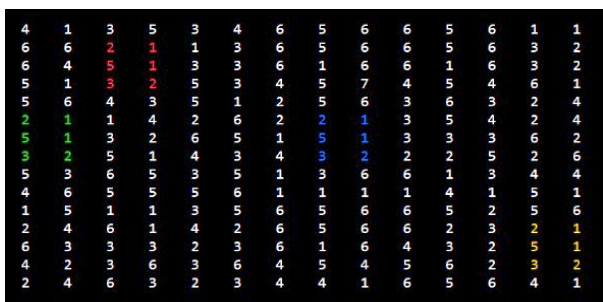


Рисунок 1 – Виділення еталонних фрагментів в цілочисельній матриці

Розроблене програмне забезпечення може застосовуватись як універсальний інструмент пошуку еталонних фрагментів у матрицях довільного обсягу. У засобі аналізу матриць передбачена можливість виділення фрагментів в таких режимах:

DECOR_DEFAULT – фрагмент не виділяється;

DECOR_UNDERLINED – ідентичні з еталоном фрагменти підкреслені (рис. 2);

DECOR_BOLD – ідентичні з еталоном фрагменти виділяються жирним шрифтом;

DECOR_COORD – визначення координат ідентичних з еталоном фрагментів;

DECOR_REVERSED – ідентичні з еталоном фрагменти мають протилежні значення кольору шрифту/фону.

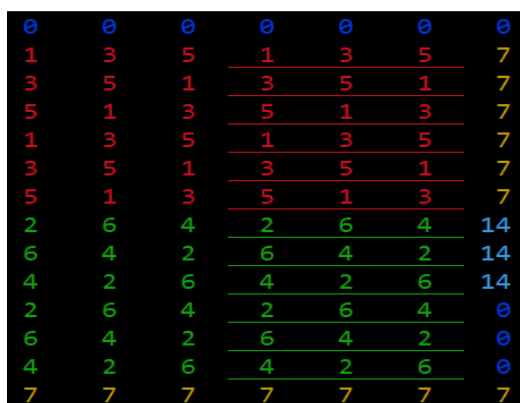


Рисунок 2 – Режим підкреслення фрагментів в процесі аналізу цілочисельної матриці

Недоліком універсального сканування є велика кількість обчислень, що зі збільшенням обсягу матриці має квадратичну залежність обчислювальної складності.

Для автоматичного синтезу алгоритмів обчислення ДКП I–IV на основі циклічних згорток необхідно забезпечити ефективний, з точки зору швидкодії, аналіз БЦС базисної матриці. Важливою особливістю структури сформованої базисної матриці є наявність набору квадратних циклічних зліва цілочисельних підматриць. Ця особливість дозволяє прискорити аналіз блочної матриці змінюючи крок сканування, що дорівнює обсягу підмасиву L_i , до якого належать елементи блоку. Програмне рішення реалізовано в IDE Visual

C++ 2019. В програмному коді на мові C++ аналіз структури матриці включає такі основні функції:

highlight_block_fragments () – виділяє підматриці в матриці (Рис. 3);

is_in_coordinates_arr() – перевіряє наявність елемента у масиві координат;

are_equal_fragments() – перевіряє на ідентичність підматриці;

coordinates is_in_arr() – пошук заданого рядка в матриці;

coordinates repeated_fragment() – пошук заданого фрагменту у матриці;

element_matrix_print() – виконує вивід проаналізованої матриці на екран.

Пошук і аналіз ідентичних підматриць виконуємо, починаючи з верхнього рядка або першого стовпця. Далі проводимо перевірку на циклічність (одинакові значення послідовності елементів зі зміщеним у наступному рядку або стовпці). При невиконанні циклічності робимо висновок про завершення обсягу підматриці, тобто можна визначити її розмірність. Далі зміщуємось (по горизонталі/по вертикалі) на цей визначений обсяг і продовжуємо аналіз перебором елементів перших рядків або стовпців і, визначаючи нову підматрицю, що ідентична попередньо виділеною. В протилежному випадку зміщення відбувається на один крок для пошуку циклічної підматриці.

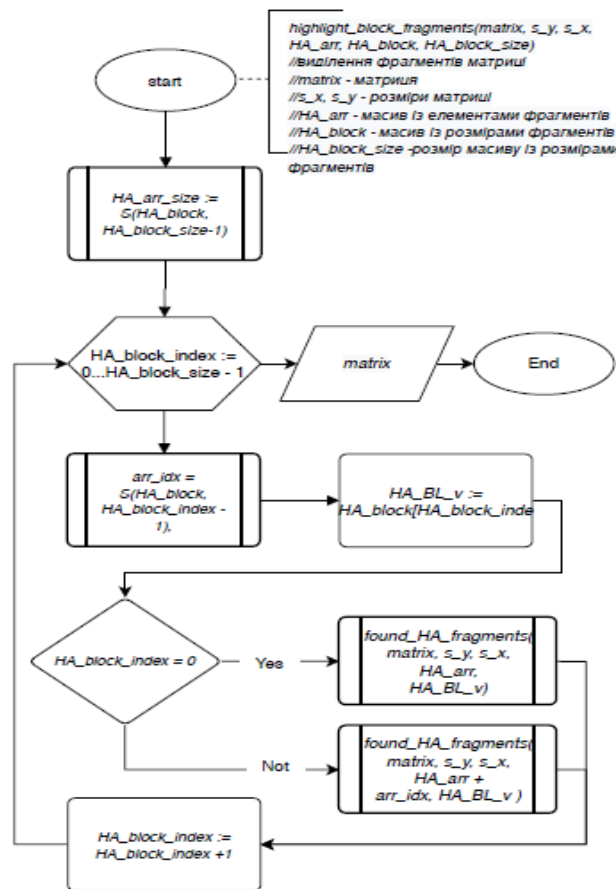


Рисунок 3 – Блок-схема функції highlight_block_fragments ()

4 ЕКСПЕРИМЕНТИ

Для проведення аналізу структури базисної матриці ДКП I–IV для послідовностей довільного обсягу N враховується, що побудовано відповідно твірному $H(L)$ блочно-циклічну базисну матрицю спрощених аргументів функцій.

Наприклад, для ДКП-II обсягу $N=20$ сформуваємо ТМ

$$H(20)=(0)(1,3,9,27)(2,6,18,26)(4,12,36,28) \\ (5,15,35,25)(7,21,17,29)(8,24)(10,30)(11,33,19,23) \\ (13,39,37,31)(14,38,34,22)(16,32)(20).$$

В результаті виконання перерозподілу твірних підмасивів, відповідно умови (5), отримаємо

$$H(20)=(0)(1,3,9,27)(39,37,31,13)(2,6,18,26) \\ (38,34,22,14)(4,12,36,28)(5,15,35,25)(7,21,17,29) \\ (33,19,23,11)(8,24)(32,16)(10,30)(20).$$

На основі твірною масиву здійснюється переіндексація рядків/стовпців спрощених аргументів функцій базисної матриці, що в результаті приводить до формування БЦС в базисній матриці. Наступним етапом синтезу алгоритмів ДКП на основі ЦЗ є виконання функції аналізу структури сформованої блочно-циклічної матриці. В результаті аналізу базисної матриці, на прикладі для ДКП-II обсягу $N=20$, отримуємо кольорове виділення циклічних підматриць та їх перших елементів у вигляді поданому на рис. 4. Відповідним кольором виділено блоки, що містять множину елементів, що належить тільки їм. Наприклад, червоним кольором (рис. 3) виділено блоки, в яких циклічні підматриці містять елементи $\{1, 3, 9, 13\}$.

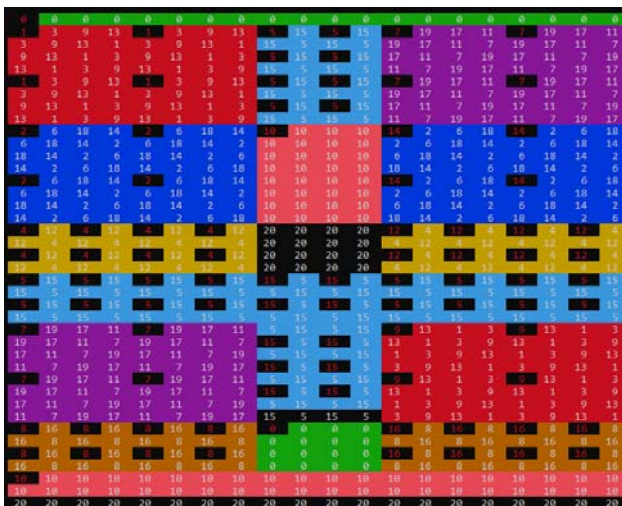


Рисунок 4 – Аналіз структури базисної матриці аргументів базисних функцій ДКП-II обсягу $N=20$

5 РЕЗУЛЬТАТИ

Проведено тестування виконання аналізу базисних матриць ДКП-II обсягу $N=20$ розробленим універсальним програмним засобом, який виконує сканування всього набору елементів матриці, та функцією аналізу

блочно-циклічної матриці зі змінним кроком сканування (рис. 5).

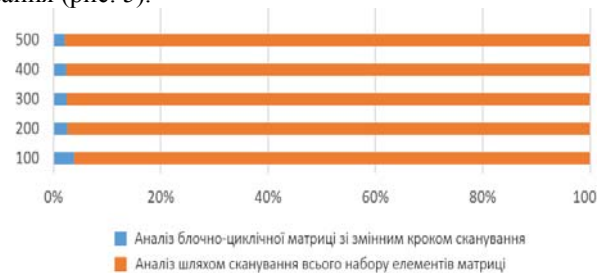


Рисунок 5 – Проведення аналізу структури базисної матриці ДКП-II обсягу $N=20$ програмами сканування з одиничним та змінним кроком

Одержано результати у відсотках часу проведення аналізу блочно-циклічної базисної матриці скануванням з одиничним та змінним кроком для кількості спроб від 100 до 500. Отже, врахування особливостей БЦС сформованої матриці аргументів базисних функцій для змінного кроку сканування, в разі прискорює процес виконання аналізу структури базисної матриці ДКП-II обсягу $N=20$.

В процесі автоматичного синтезу алгоритмів обчислення ДКП I–IV довільного обсягу N на основі ЦЗ завершенням етапу аналізу БЦС базисної матриці є формування масиву даних, що інформує про кількість ідентичних циклічних підмасивів, їх розташування. Ці дані дозволяють визначити мінімальну кількість та обсяг циклічних згортки, можливість їх паралельного обчислення та об'єднання результатів ЦЗ для визначення вихідних значень ДКП.

В результаті, для визначення ідентичних блоків, розміщених горизонтально та вертикально один відносно іншого, функція формує масив перших елементів циклічних підмасивів з відповідними координатами. Наприклад, для ДКП-II обсягу $N=20$ отримуємо масив даних

- 1{1,0}; 1{5,0}; 1{1,4}; 1{5,4}; 9{25,12}; 9{29,12}; 9{25,16}; 9{29,16};
- 2{9,0}; 2{13,0}; 2{9,4}; 2{13,4}; 14{9,12}; 14{13,12}; 14{9,16}; 14{13,16};
- 4{17,0}; 4{19,0}; 4{17,2}; 4{19,2}; 4{17,4}; 4{19,4}; 4{17,6}; 4{19,6}; 12{17,12}; 12{19,12}; 12{17,14}; 12{19,14}; 12{17,16}; 12{19,16}; 12{17,18}; 12{19,18};
- 5{21,0}; 5{23,0}; 5{21,2}; 5{23,2}; 5{21,4}; 5{23,4}; 5{21,6}; 5{23,6}; 5{1,8}; 5{3,8}; 5{5,8}; 5{7,8}; 5{1,10}; 5{3,10}; 5{5,10}; 5{7,10}; 5{21,12}; 5{23,12}; 5{21,14}; 5{23,14}; 5{21,16}; 5{23,16}; 5{21,18}; 5{23,18}; 15{21,8}; 15{23,8}; 15{26,8}; 15{28,8}; 15{30,8}; 15{21,10}; 15{23,10}; 15{26,10}; 15{28,10}; 15{30,10}; 7{25,0}; 7{29,0}; 7{25,4}; 7{29,4}; 7{1,12}; 7{5,12}; 7{1,16}; 7{5,16};
- 8{33,0}; 8{35,0}; 8{33,2}; 8{35,2}; 8{33,4}; 8{35,4}; 8{33,6}; 8{35,6}; 16{33,12}; 16{35,12}; 16{33,14}; 16{35,14}; 16{33,16}; 16{35,16}; 16{33,18}; 16{35,18};
- 0{0,0}; 0{33, 8}; 10{9,8}; 10{37,0}.

Відповідно виконаному аналізу базисної матриці на рис. 3 перші елементи циклічних підмасивів, що дорівнюють, наприклад, 1 розташовані в базисній матриці за координатами $1\{1,0\}$; $1\{5,0\}$; $1\{1,4\}$; $1\{5,4\}$.

Тобто, маємо в структурі базисної матриці по дві ідентичні підматриці розміщених горизонтально $1\{1,0\}$ і $1\{1,4\}$ та $1\{5,0\}$ і $1\{5,4\}$, вертикально $1\{1,0\}$ і $1\{5,0\}$ та $1\{1,4\}$ і $1\{5,4\}$ одна відносно іншої.

Масив перших елементів з координатами розташування циклічних підмасивів в максимальному випадку може мати $(n_1 \times n_2)$ елементів в структурі базисної матриці $V_N(n_1, n_2)$ спрощених аргументів ДКП I–VI довільного обсягу N .

6 ОБГОВОРЕННЯ

Наявність ідентичних підматриць в структурі базисної матриці ДКП приводить до зменшення обчислювальної складності. Обсяг перетворення N та вид ДКП визначають кількість циклічних підматриць в структурі базисної матриці [12]. Ідентичні циклічні підматриці, що розміщені вертикально один відносно іншого приводять до одноразового обчислення ЦЗ, результати якої в процесі об'єднання використовуються для різних $X^c(m)$ вихідних значень перетворення. Ідентичні циклічні підматриці, що розміщені горизонтально один відносно іншого приводять до об'єднання груп вхідних значень $x(n)$ перетворення і одноразового обчислення ЦЗ, результати якої в процесі об'єднання використовуються для однієї групи вихідних значень перетворення. Так, дані аналізу $1\{1,0\}$; $1\{5,0\}$; $1\{1,4\}$; $1\{5,4\}$ в структурі базисної матриці (рис. 3) вказують на наявність двох ідентичних підматриць розміщених горизонтально та двох ідентичних підматриць розміщених вертикально одна відносно іншої. В цьому випадку виконується спершу об'єднання груп вхідних значень для підматриць розміщених горизонтально виконується одноразове обчислення ЦЗ, результати якої в процесі об'єднання використовуються для різних $X^c(m)$ вихідних значень перетворення, залежних від ідентичних підматриць розміщених вертикально.

Визначення мінімальної кількості ЦЗ на основі параметрів, які характеризують БЦС базисної матриці ДКП, надає можливість їхнього розпаралелення обчислення і тим самим до пришвидшення виконання перетворення.

Таким чином, формування базисної матриці ДКП у вигляді набору циклічних підматриць та проведення аналізу блочної структури дозволяє в результаті забезпечити високу ефективність використання обчислювальних ресурсів, та зменшити обчислювальну складність, що залежить як від обсягу так і виду ДКП.

ВИСНОВКИ

У роботі розглянуто аналіз на ідентичність підматриць зі змінним кроком пошуку, що враховує властивості блочності та циклічності сформованої базисної матриці ДКП I–VI. Розроблено алгоритмічне та програмне забезпечення виконання аналізу структури

© Процько І. О., Мішук М. В., 2020
DOI 10.15588/1607-3274-2020-3-16

блочно-циклічної базисної матриці, за допомогою якого визначається масив параметрів даних формального опису структури базису ДКП. Аналіз БЦС базисної матриці, що в результаті формує масив даних про кількість вертикально та горизонтально розташованих ідентичних циклічних підматриць, є важливим етапом автоматичного синтезу алгоритмів обчислення ДКП I–IV на основі ЦЗ довільних обсягів N .

Наукова новизна полягає у визначенні особливостей аналізу та обчислення параметрів, які характеризують структуру блочно-циклічної базисної матриці ДКП.

Практичне значення розробленого алгоритмічного та програмного забезпечення виконання аналізу БЦС базисної матриці ДКП полягає у забезпеченні зменшення обчислювальної складності та можливості розпаралелення обчислення синтезованого швидкого алгоритму обчислення ДКП.

Напрямок подальших досліджень полягатиме в розробці алгоритмічного та програмного забезпечення що зменшуватиме обчислювальні затрати та прискорить процес синтезу швидкого алгоритму обчислення ДКП, включаючи й етап аналізу БЦС базисної матриці перетворення.

ПОДЯКИ

Робота виконана в рамках держбюджетної науково-дослідної роботи ДБ / Нейрозахист (номер держ. реєстрації № ДР 0119U002256) (2019–20р.) національного університету «Львівська політехніка»

ЛІТЕРАТУРА / LITERATURA

1. Horn R. A. Matrix analysis / R. A. Horn, C. R. Johnson. – New York: Cambridge university press, 1985. – 561p. DOI: 10.1017/CBO9780511810817
2. Trott M. The Mathematica GuideBook for Programming / M. Trott. – New York : Springer-Verlag, 2004. – 1028 p. DOI 10.1007/978-1-4419-8503-3
3. Chan Y.-H. Generalized approach for the realization of discrete cosine transform using cyclic convolutions / Y.-H. Chan, W.-C. Siu // IEEE international conference on Acoustics, Speech, and Signal processing: digital speech processing, Minneapolis, USA, 27–30 April 1993: proceedings. – Washington: IEEE Computer Society, DC, 1993. – Vol. III. – P. 277–280. DOI: 10.1109/ICASSP.1993.319489
4. Prots'ko I. Algorithm of Efficient Computation of DCT I–IV Using Cyclic Convolutions / I. Prots'ko // International Journal of Circuits, Systems and Signal Processing. – 2013. – Vol. 7, Issue 1. – P. 1–9.
5. Prots'ko I. Performance evaluation of the program of DCT-II using cyclic convolutions / I. Prots'ko, R. Rikmas, M. Mashevska // Computer Sciences and Information Technologies (CSIT'2017): International scientific and technical conference, Lviv, Ukraine, 5–8 September 2017: proceedings. – Lviv: Lvivska polytechnika Press, 2017. – P. 276–278. DOI: 10.1109/STC-CSIT.2017.8098785
6. Процько І. О. Особливості обчислення твірних масивів для синтезу швидких алгоритмів ДКП I–IV / І. О. Процько // Радіоелектроніка, інформатика, управління. – 2020. – № 2. – P. 149–157.
7. Image processing and matrices / [O. Prabhune, P. Sabale, D. N. Sonawane, C. L. Prabhune] // Data Management, Ana-

- lytics and Innovation (ICDMAI): International conference, Pune, India, 24–26 Feb. 2017: proceedings. – Pune : Curran Associates, Inc., 2017. – P. 166–171. DOI: 10.1109/ICDMAI.2017.8073504
8. Performance evaluation and analysis of sparse matrix and graph kernels on heterogeneous processors / [F. Zhang, W. Liu, N. Feng et al.] // CCF Transactions on High Performance Computing. – 2019. – Vol. 1. – P. 131–143.
9. Duff I. S. An overview of the sparse basic linear algebra subprograms: the new standard from the BLAS technical forum / I. S. Duff, M. A. Heroux, R. Pozo // ACM Transactions on Mathematical Software. – 2002. – Vol. 28, No. 2. – P. 239–267.
10. Pichel J. C. Sparse Matrix Classification on Imbalanced Datasets Using Convolutional Neural Networks / J. C. Pichel, B. Pateiro-Lopez // IEEE Access. – 2019. – Vol. 7. – P. 82377–82389. DOI: 10.1109/ACCESS.2019.2924060
11. Prots'ko I. The Algorithm and Structures for Efficient Computation of Type II/III DCT/ DST/ DHT Using Cyclic Convolutions / I. Prots'ko // International Journal of Signal Processing Systems. – 2014. – Vol. 2, No. 2. – P. 119–127. DOI: 10.12720/ijsp.2.2.119–127
12. Prots'ko I. Repeatability the block cyclic structures the basis matrices of DCT for sizes p^n / I. Prots'ko, R. Rykmas, V. Teslyuk // International conference Perspective Technologies and Methods in MEMS Design (MEMSTECH'2015), Polyana-Svalyava, Ukraine, 2–4 September 2015: proceedings. – Lviv : Veza & Co, 2015. – P. 107–109.

Стаття надійшла до редакції 25.05.2020.
Після доробки 01.06.2020.

УДК 004.93

ПРОГРАММНЫЙ АНАЛИЗ СТРУКТУРЫ БЛОЧНО-ЦИКЛИЧЕСКОЙ БАЗИСНОЙ МАТРИЦЫ ДКП

Процько І. О. – д-р техн. наук, доцент кафедри автоматизованих систем управління Національного університету «Львівська політехніка», Україна.

Мищук М. В. – студент кафедри автоматизованих систем управління Національного університету «Львівська політехніка», Україна.

АННОТАЦИЯ

Актуальность. Использование матричного записи применяется для формализации предметной области в рамках алгебраического подхода. Эффективное вычисления дискретного косинусного преобразований использует приведение гармоничного базиса к блочно-циклической матричной структуре с последующим вычислением преобразования с помощью быстрых циклических сверток. Анализ структуры базисной блочной матрицы преобразования обеспечивает синтез эффективных алгоритмов дискретного косинусного преобразования произвольных размеров. Программная реализация проведения анализа за блочно-циклических структур формирует описание структуры, что позволяет уменьшить вычислительную сложность алгоритма преобразования и выполнить распараллеливания вычислений циклических сверток.

Цель работы – определение алгоритмических особенностей проведения анализа структуры блочно-циклической матрицы, содержащей целочисленные аргументы базисных гармонических функций, что обеспечит уменьшение вычислительной сложности синтезированного алгоритма дискретного косинусного преобразования на основе циклических сверток.

Метод. Поиск и анализ перебором элементов матрицы с переменным шагом, учитывающий свойства блочности и цикличности сформированной базисной матрицы дискретного косинусного преобразования, позволяет быстрее выполнять анализ структуры блочной матрицы преобразования по сравнению с полным сканированием.

Результаты. Разработано алгоритмическое и программное обеспечение выполнения анализа структуры блочно-циклической базовой матрицы, с помощью которого определяется массив параметров данных формального описания структуры базиса дискретного косинусного преобразования. Данные анализа структуры базовой матрицы позволяют определить наличие идентичных циклических подматриц размещенных горизонтально или вертикально друг относительно друга и, тем самым, уменьшить количество выполнения циклических сверток.

Выводы. Эффективный анализ блочно-циклической структуры базисной матрицы на основе разработанного программного обеспечения является важной частью процесса синтеза быстрого алгоритма, который обеспечивает сокращение вычислительной сложности и возможность распараллеливания выполнения дискретного косинусного преобразования. Разработанное алгоритмическое и программное обеспечение выполнения анализа структуры блочно-циклической матрицы может применяться и для анализа структуры и поиска соответствующих подматриц в любых матрицах с целыми, действительными и нулевыми элементами.

КЛЮЧЕВЫЕ СЛОВА: анализ матрицы, алгоритм поиска, образующий массив, блочно-циклические подматрицы, дискретные косинусные преобразования.

UDC 004.93

SOFTWARE ANALYSIS OF STRUCTURE BLOCK-CYCLIC BASIC MATRIX OF DCT

Protsko I. O. – Dr. Sc., Associate Professor, Department of Automated Control Systems, Lviv Polytechnic National University, Lviv, Ukraine.

Mishchuk M. V. – Student, Department of Automated Control Systems, Lviv Polytechnic National University, Lviv, Ukraine.

ABSTRACT

Context. The matrix notation is used to formalize the subject area within the framework of the algebraic approach. Effective computation of the discrete cosine transforms uses the reduction of a harmonic basis to a block-cyclic matrix structure with the subsequent calculation of the transform using fast cyclic convolutions. An analysis of the structure of the basic block matrix of transforms provides a synthesis of algorithms of effective discrete cosine transforms of arbitrary sizes. The software implementation of the analysis of block-cyclic structures generates a description of the structure, which allows to reduce the computational complexity of the algorithm of effective discrete cosine transform and to perform parallelization of computation the cyclic convolutions.

Objective. The work is to determine the algorithmic features of the analysis of the structure of a block-cyclic matrix containing integer arguments of basic harmonic functions, which will reduce the computational complexity of the synthesized discrete cosine transform algorithm based on cyclic convolutions.

Method. Search and analysis by enumerating elements of the matrix with a variable step, taking into account the blockiness and cyclicity of the formed basis matrix of the discrete cosine transform, allows you to quickly analyze the structure of the block matrix of transform in comparison with full scanning.

Results. Algorithmic and software for analyzing the structure of a block-cyclic basis matrix have been developed, with the help of which an array of data parameters for a formal description of the basis matrix structure of a discrete cosine transform is determined. The analysis of the structure of the base matrix allows us to determine the presence of identical cyclic submatrices placed horizontally or vertically relative to each other and, thereby, reduce the number of cycles of convolutions.

Conclusions. An effective analysis of the block-cyclic structure of the basis matrix based on the developed software is an important part of the fast algorithm synthesis process, which provides a reduction in computational complexity and the ability to parallelize the implementation of the discrete cosine transform. The developed algorithmic and software for performing the analysis of the structure of a block-cyclic matrix can also be used to analyze the structure and search for the corresponding submatrices in any matrices with integer, real, and zero elements.

KEYWORDS: matrix analysis, search algorithm, hashing array, block-cyclic submatrices, discrete cosine transforms.

REFERENCES

1. Horn R. A., Johnson C. R. Matrix analysis. New York, Cambridge university press, 1985, 561 p. DOI: 10.1017/CBO9780511810817
2. Trott M. The Mathematica GuideBook for Programming. New York, Springer-Verlag, 2004, 1028 p. DOI 10.1007/978-1-4419-8503-3
3. Chan Y.-H., Siu W.-C. Generalized approach for the realization of discrete cosine transform using cyclic convolutions, *IEEE international conference on Acoustics, Speech, and Signal processing: digital speech processing, Minneapolis, USA, 27–30 April 1993: proceedings*. Washington, IEEE Computer Society, DC, 1993, Vol. III, pp. 277–280. DOI: 10.1109/ICASSP.1993.319489
4. Prots'ko I. Algorithm of Efficient Computation of DCT I–IV Using Cyclic Convolutions, *International Journal of Circuits, Systems and Signal Processing*, 2013, Vol. 7, Issue 1, pp. 1–9.
5. Prots'ko I., Rikmas R., Mashevskaya M. Performance evaluation of the program of DCT-II using cyclic convolutions, *Computer Sciences and Information Technologies (CSIT'2017): International scientific and technical conference, Lviv, Ukraine, 5–8 September 2017: proceedings*. Lviv, Lvivska polytechnika Press, 2017, pp. 276–278. DOI: 10.1109/STC-CSIT.2017.8098785
6. Prots'ko I. O. Peculiarities of computation the hashing arrays for the synthesis of fast algorithms of DCT I–IV, *Radio Electronics, Computer Science, Control*, 2020, No. 2, pp. 149–157.
7. Prabhune O., Sabale P., Sonawane D. N., Prabhune C. L. Image processing and matrices, *Data Management, Analytics and Innovation (ICDMAI): International conference, Pune, India, 24–26 Feb. 2017: proceedings*. Pune, Curran Associates, Inc., 2017, pp. 166–171. DOI: 10.1109/ICDMAI.2017.8073504
8. Zhang F., Liu W., Feng N., Zhai J., Du X. Performance evaluation and analysis of sparse matrix and graph kernels on heterogeneous processors, *CCF Transactions on High Performance Computing*, 2019, Vol. 1, pp. 131–143.
9. Duff I. S., Heroux M. A., Pozo R. An overview of the sparse basic linear algebra subprograms: the new standard from the BLAS technical forum, *ACM Transactions on Mathematical Software*, 2002, Vol. 28, No. 2, pp. 239–267.
10. Pichel J. C., Pateiro- B. Lopez Sparse Matrix Classification on Imbalanced Datasets Using Convolutional Neural Networks, *IEEE Access*, 2019, Vol. 7, pp. 82377–82389. DOI: 10.1109/ACCESS.2019.2924060
11. Prots'ko I. The Algorithm and Structures for Efficient Computation of Type II/III DCT/ DST/ DHT Using Cyclic Convolutions, *International Journal of Signal Processing Systems*, 2014, Vol. 2, No. 2, pp. 119–127. DOI: 10.12720/ijspss.2.2.119–127
12. Prots'ko I., Rykmas R., Teslyuk V. Repeatability the block cyclic structures the basis matrices of DCT for sizes p^n , *International conference Perspective Technologies and Methods in MEMS Design (MEMSTECH'2015), Polyana-Svalyava, Ukraine, 2–4 September 2015: proceedings*. Lviv, Veza & Co, 2015, pp. 107–109.

МОДЕЛИ ВЫЧИСЛЕНИЯ ВЕСОВ ДЛЯ ОЦЕНКИ ИННОВАЦИОННЫХ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

Титенко Е. А. – канд. техн. наук, доцент, ведущий научный сотрудник центра перспективных исследований и разработок, ФГБОУ ВО «Юго-Западный государственный университет», г. Курск, Россия.

Фролов С. Н. – канд. техн. наук, ведущий научный сотрудник центра перспективных исследований и разработок, ФГБОУ ВО «Юго-Западный государственный университет», г. Курск, Россия.

Ханис А. Л. – канд. воен. наук, доцент кафедры информационной безопасности, ФГБОУ ВО «Юго-Западный государственный университет», г. Курск, Россия.

Добросердов О. Г. – д-р техн. наук, старший научный сотрудник, советник при ректорате, ФГБОУ ВО «Юго-Западный государственный университет», г. Курск, Россия.

Захаренков А. И. – д-р техн. наук, профессор, первый заместитель генерального директора, АО «Воентелеком», г. Москва, Россия.

Попов А. Н. – генеральный директор, АО «Авиаавтоматика» им. В. В. Тарасова», г. Курск, Россия.

Дронова Т. А. – д-р мед. наук, профессор кафедры пропедевтики внутренних болезней, ФГБОУ ВО «Курский государственный медицинский университет» Министерства здравоохранения Российской Федерации, г. Курск, Россия.

АННОТАЦИЯ

Актуальность. Актуальность работы связана с многокритериальным сравнением инновационных объектов, под которыми понимаются объекты с частично совпадающими индивидуальными показателями, имеющие неопределенные, субъективные значения и критерии их оценки. Сравнимые инновационные объекты описываются общими и индивидуальными показателями, причем количество индивидуальных показателей является значительным (более 50%). Метод анализа иерархий является наиболее подходящим математическим аппаратом для сравнения инновационных объектов, так как он совмещает наглядность описания предметной области, численную результативность, естественный учет вариативности исходных парных оценок. Тем не менее, данный метод слабо учитывает индивидуальные показатели инновационных объектов при составлении матриц парных сравнений, что не позволяет получить итоговые веса, соответствующие индивидуальным особенностям инновационных объектов.

Цель. Разработка модифицированного метода анализа иерархий, обеспечивающего ранжирование инновационных объектов с частично совпадающими индивидуальными показателями.

Метод. В работе разработан модифицированный метод анализа иерархий, включающий типовую последовательность этапов от построения иерархии предметной области и формирования критериев оценки инновационных объектов до этапа вычисления итоговых весов инновационных объектов, отличающийся во-первых, введением на этапе построения иерархии предметной области матрицы весов общих и индивидуальных показателей по системам объектов, во-вторых расчётом поправочных коэффициентов на основе локальной или глобальной моделей коррекции весов показателей систем объектов на этапе составления сводной таблицы весов, в-третьих вычислением итоговых весов инновационных объектов с учётом поправочных коэффициентов, что позволило учесть индивидуальные особенности сравниваемых инновационных объектов.

Результаты. В работе рассмотрены инновационные объекты с общими и индивидуальными показателями в пропорции 1:2 с количеством критериев сравнения, равным 5, что соответствует классу объектов средней организационной сложности. Сравнимые объекты содержат типовой объект (количество индивидуальных показателей существенно меньше 50%), явно инновационный объект (количество индивидуальных показателей больше 50%), объект в пограничной зоне (количество индивидуальных показателей около 50%). Классический метод не чувствителен к индивидуальным особенностям инновационных объектов, что определило минимальный вес для явно инновационного объекта. Модифицированный метод, напротив, определил минимальный вес для объекта в пограничной зоне, так как он имеет одновременно пониженные значения в матрице парных сравнений по критериям и низкие значения индивидуальных показателей.

Выводы. Разработанный модифицированный метод анализа иерархий и модели коррекции в его составе объективно отражают порядок ранжирования объектов с учетом их описания в виде матрицы общих и индивидуальных показателей. Согласно моделированию, ценным для практики принятия решений следует считать долю индивидуальных показателей в их общем количестве более 55 %. В этом случае объекты предпочтительнее оценивать на основе модифицированного метода анализа иерархий. Полученный на основе модели коррекции весов вектор поправочных коэффициентов имеет самостоятельное значение при решении различных расчетно-аналитических задач и прикладных задач принятия решений.

КЛЮЧЕВЫЕ СЛОВА: инновационный высокотехнологичный объект, модель коррекции весов, расширенная сводная таблица, вектор поправочных коэффициентов, метод анализа иерархий.

АББРЕВИАТУРЫ

MA – среднеарифметическое значение;

MGM – среднегеометрическое значение;

MHR – среднегармоническое значение;

БЛА – беспилотные летательные аппараты;

МАИ – метод анализа иерархий;

МПС – матрицы парных сравнений.

НОМЕНКЛАТУРА

A – матрица;

$Arg(i)$ – целевая функция выбора объектов;
 C – управляющая система;
 d – среднее количество индивидуальных показателей на систему в составе объекта;
 E – энергетическая система;
 z – количество критериев сравнения;
 K_x – x -й критерий сравнения объекта $x=1-z$;
 K_1 – критерий надежности;
 K_2 – критерий ремонтпригодности;
 K_3 – критерий комплектации;
 K_4 – критерий стоимости;
 K_5 – критерий летательных качеств;
 L – связанная система;
 m – количество сравниваемых объектов (альтернатив);
 N – навигационная система;
 O_{ij} – вес i -го объекта по j -му критерию ($i=1-m$, $j=1-q$);
 P – общее количество индивидуальных показателей;
 p – количество индивидуальных показателей системы;
 p_i^* – подмножество индивидуальных показателей;
 q – количество систем;
 S – общее количество общих показателей;
 s – количество общих показателей системы;
 s_i^* – подмножество общих показателей;
 T – движительная система;
 VCC_{MA} – вектор поправочных коэффициентов на основе среднеарифметического;
 VCC_{MGM} – вектор поправочных коэффициентов на основе среднегеометрического;
 vs_{ij} – вес ij -го общего показателя;
 vp_{ij} – вес общего показателя i -го объекта для j -й системы;
 vs_{ij}' – скорректированный вес общего ij -го показателя;
 w_i – i -й элемент вектора весов ($i=1-q$);
 θ – функция вычисления веса i -го объекта;
 φ – сборочная функция.

ВВЕДЕНИЕ

Задачи анализа и многокритериального выбора альтернатив для инновационных объектов связаны с обработкой неопределенной, неточной, субъективной информации о характеристиках объектов и критериях оценки. Типовые примеры таких задач:

- сравнение сложных технических комплексов, технологий/методик производства наукоемкой продукции;
- выбор типа подвижного робота, беспилотного летательного аппарата;
- оценка инновационных проектов;
- поиск перспективной научно-технической продукции;
- оценка рисков производства высокотехнологичной продукции и др.

Традиционно сравнение и оценка объектов в рамках теории принятия решений [1, 2] основаны на:

© Титенко Е. А., Фролов С. Н., Ханис А. Л., Добросердов О. Г., Захаренков А. И., Попов А. Н., Дронова Т. А., 2020
DOI 10.15588/1607-3274-2020-3-17

- сложившихся весовых отношениях между показателями;
- общезначимых для предметной области критериях сравнения;
- объективных, точных оценках сравниваемых альтернатив;
- унификации множества показателей, описывающих объекты сравнения.

К настоящему времени разработано множество моделей и методов решения задачи многокритериального выбора – математические методы сверток, модели интегральных показателей, модели нечетких оценок, методы структурного моделирования, иерархического описания [3–12] и др. Наибольший интерес представляют такие модели и методы, которые дают возможность учитывать многокритериальность, неопределенность исходных данных, субъективизм промежуточных оценок. Наиболее остро данные ограничения проявляются при сравнении инновационных объектов (альтернатив), так как они характеризуются субъективными значениями показателей (новый/известный, выгодный/убыточный, привлекательный/традиционный, автономный/зависимый и др.).

Классический МАИ, предложенный Т. Л. Саати [13, 14], и его многочисленные модификации наиболее полно подходят для сравнения и выбора инновационных объектов [15–20]. МАИ совмещает наглядность описания предметной области, численную результативность, естественный учет вариативности исходных данных.

Как известно, главная идея данного метода состоит в парном сопоставлении критериев, построении МПС и их обработке с целью составления сводной матрицы весов объектов по всем критериям, вычисления по ней итоговых весов объектов и выбора лучшей альтернативы. Достаточно полная классификация модификаций МАИ представлена в работах Мироновой Н. А. [16, 17]. В них систематизированы модификации МАИ, применимые для различных типов данных, шкал измерений, типа иерархии, вида системы предпочтения экспертов, метода формирования МПС, применяемых правил свертки и др.

Тем не менее, МАИ при сравнении инновационных объектов слабо учитывает их индивидуальные показатели, что не всегда позволяет получить объективные парные оценки альтернатив или полностью заполнить все МПС [21, 22]. Эта ситуация составляет основное противоречие исследования.

В связи с этим в работе предлагается подход к уточнению МПС с учетом структурно-функциональной организации инновационных объектов. Сравнимые инновационные объекты по своим составным частям (системам) описываются общими и индивидуальными показателями, причем количество индивидуальных показателей является значительным (более 50%). Данная особенность наиболее характерна для высокотехнологичных инновационных объектов: робототехнические комплексы, БЛА, малые космические аппараты, коммуникационные ИТ-

устройства, материалы с метеорологическими меняющимися свойствами, экзоскелеты, станки с изменяемой операционной частью, двигатели-гибриды и др.

Объектом исследований является сравнительный анализ инновационных высокотехнологичных технических объектов, таких как БЛА, малые космические аппараты, наземные робототехнические комплексы, обладающих иерархичной структурой, слабой унификацией.

Предметом исследований является модифицированный МАИ, использующий модели вычисления весов инновационных объектов по наборам общих и индивидуальных показателей и позволяющий осуществить выбор лучшей альтернативы при субъективных оценках экспертов.

Целью работы является разработка модифицированного МАИ, обеспечивающего ранжирование инновационных технических объектов с частично совпадающими индивидуальными показателями.

1 ПОСТАНОВКА ЗАДАЧИ

В работе в качестве примера инновационных технических объектов рассматриваются БЛА. Они характеризуются большим набором разнородных показателей частично присутствующих в различных экземплярах БЛА [23–26].

Пусть для повышения обоснованности принимаемых решений БЛА [23–25] декомпозируется на q систем ($q = 5$) – T, C, N, E, L . Каждая из этих систем описывается множествами общих и индивидуальных показателей S и P соответственно. Тогда вычисление приоритетов (весов) по m сравниваемым объектам (альтернативам) $O_1 \dots O_m$ и выбор лучшей альтернативы осуществляется на основе целевой функции вида

$$\begin{aligned} & Arg(i) | \theta(O_i, K_1 \dots K_z, T(s_i, p_i), C(s_i, p_i), \\ & N(s_i, p_i), E(s_i, p_i), L(s_i, p_i)) \rightarrow \max. \end{aligned} \quad (1)$$

2 ОБЗОР ЛИТЕРАТУРЫ

Метод анализа иерархий разработан в 1970 году американским ученым Томасом Л. Саати. В настоящее время он активно развивается и широко используется в практической деятельности для решения самых разнообразных задач. Метод позволяет лицу, принимающему решение, структурировать сложную проблему в виде иерархии и выполнить количественную оценку имеющихся вариантов решения (альтернатив). Результаты практического применения данной методологии приведены во многих российских и зарубежных изданиях, в частности, существуют издания, целиком посвященные развитию МАИ [13–20].

Метод анализа иерархий постоянно расширялся и модифицировался. Изменения были направлены на повышение обоснованности принимаемых решений, повышение точности решения в задаче вычисления собственного вектора матрицы, уменьшение вычислительной сложности проверочных шагов, введение

нелинейных решающих функций на заключительных этапах метода [15–22]. Тем не менее, этапы формирования и проверки матриц парных сравнений в части уточнения значений этих матриц рассматривались без привязки к структурно-функциональной организации инновационных объектов. Узким местом метода является несогласованность оценок экспертов при сравнении инновационных объектов в силу наличия нефакторов у объектов [21], неоднозначности ранжирования критериев [22] при итоговой свертке. Это обстоятельство делало метод существенно зависящим от мнения и квалификации эксперта, т.е. вносило значительный субъективизм на этапе ввода исходных данных – МПС [16–18, 20].

3 МАТЕРИАЛЫ И МЕТОДЫ

Как известно [4, 9], классический МАИ состоит из последовательности этапов:

- 1) построение иерархии предметной области и формирование критериев оценки объектов;
- 2) составление матрицы парных сравнений критериев;
- 3) получение вектора приоритетов критериев;
- 4) оценка степени согласованности матрицы парных сравнений критериев;
- 5) составление матриц парных сравнений сравниваемых объектов по каждому критерию;
- 6) получение векторов приоритетов объектов по каждому критерию;
- 7) оценка степени согласованности всех МПС;
- 8) составление сводной матрицы весов объектов по всем критериям;
- 9) вычисление глобального вектора приоритетов объектов и выбор лучшей альтернативы.

Применительно к инновационным объектам особенность первого этапа МАИ заключается в том, что такие объекты декомпозируются на q систем, которые в дальнейшем описываются матрицами общих и индивидуальных показателей. Сравнение объектов осуществляется по K критериям. При этом справедливо, что

$$s \in T, s \in C, s \in L, s \in E, s \in N, \quad (2)$$

$$\forall(i) (s_i \neq 0), i=1-q. \quad (3)$$

Для индивидуальных показателей в соответствии с $p = p_1 \cup p_2 \cup p_3 \cup p_4 \cup p_5$ будет справедливо, что

$$p_1 \in T, p_2 \in C, p_3 \in L, p_4 \in E, p_5 \in N, \quad (4)$$

$$\exists(i) | (p_i=0), i=1-q. \quad (5)$$

Эта особенность не позволяет напрямую использовать МАИ и его многочисленные модификации, так как каждый объект имеет многомерное описание, что затрудняет заполнение матриц парных сравнений [23–26].

Для снятия данного ограничения в рамках первого этапа МАИ предлагается формализовать наборы общих и индивидуальных показателей. Формализация сводится к созданию локальной и глобальной моделей коррекции общих показателей по сравниваемым объектам и формированию матрицы скорректированных весов.

В дальнейшем на этапе составления сводной матрицы весов объектов полученная скорректированная матрица весов является основой для вычисления вектора поправочных коэффициентов по сравниваемым объектам. Этот вектор как дополнительный сомножитель используется при вычислении итоговых весов в сводной таблице.

Таким образом, модификация МАИ включает расширение первого и предпоследнего этапов для получения и обработки общих и индивидуальных показателей по сравниваемым инновационным объектам, что позволяет учесть их индивидуальные особенности.

В работе используется количественный подход к оценке вклада показателей систем в оценку БЛА (инновационного технического объекта), а именно:

- чем больше количество показателей системы, тем больший вклад соответствующая система вносит в оценку объекта;

- чем больше количество индивидуальных показателей системы, тем больше должен быть вес показателя этой системы.

Вклад каждого из показателей систем в оценку БЛА оценивается локальным весом показателя в шкале [0–1]. Значение веса системы определяется операцией коррекции весов показателей от q систем со стороны значений индивидуальных показателей, смещающих в большую сторону вес общего показателя в пределах каждой системы.

Согласно количественному подходу подмножеству общих показателей s_i^* , принадлежащих i -му объекту ($i=1-m$), соотносится с количеством s , а подмножество индивидуальных показателей p_i^* , принадлежащих i -му объекту ($i=1-m$), соотносится со средним количеством индивидуальных показателей на систему d . При этом особенность учета вклада индивидуальных показателей определяется тем, что индивидуальный показатель отдельной системы в составе объекта предлагается дополнительно соотносить с количеством всех индивидуальных показателей для этой системы по m объектам к максимально возможному количеству индивидуальных показателей по m объектам. Эта особенность позволяет учитывать разброс количества индивидуальных показателей по сравниваемым объектам.

Значения весов общих и индивидуальных показателей по j -й системе ($j=1-q$) определяются как

$$\forall j | (vs_{ij} = s_{ij}^* / s), j=1-q, \quad (6)$$

$$\forall j | (vp_{ij} = (p_{ij}^* / d) \cdot (\sum_{i=1}^m p_{ij}^* / m \cdot d)), j=1. \quad (7)$$

Формируемая матрица весов общих и индивидуальных показателей (табл. 1) представляет собой дополнительную информацию о сравниваемых объектах $O_1 - O_m$, которую нужно ввести в МАИ в виде поправочных коэффициентов.

Таблица 1 – Матрица весов общих и индивидуальных показателей объектов по q системам

Objects	1		2		...	q	
	s	p ₁	s	p ₂		s	p _m
O ₁	vs ₁₁	vp ₁₁	vs ₁₂	vp ₁₂	...	vs _{1q}	vp _{1q}
O ₂	vs ₂₁	vp ₂₁	vs ₂₂	vp ₂₂	...	vs _{2q}	vp _{2q}
...
O _m	vs _{m1}	vp _{m1}	vs _{m2}	vp _{m2}	...	vs _{mq}	vp _{mq}

Для учета влияния индивидуальных показателей систем используются локальная или глобальная стимулирующие модели, смещающие вес общих показателей в большую сторону.

Локальная стимулирующая модель основана на персональном изменении веса общего показателя в пределах соответствующей системы [26]. Данная модель имеет вид:

$$vs'_{ij} = vs_{ij} (1 + vp_{ij}). \quad (8)$$

Свертка в итоговое значение всех индивидуальных показателей ($p_{11}, p_{12}, \dots, p_{mq}$) является сущностью глобальной стимулирующей модели вклада показателей систем инновационного объекта. Выделяются две разновидности глобальной стимулирующей модели:

- глобальная модель, основанная на равноправном вкладе всех индивидуальных показателей в итоговую оценку объекта;

- глобальная модель, основанная на учете индивидуальных показателей в пределах одного объекта сравнения.

В первом случае коррекция веса общего показателя вычисляется как:

$$vs'_{ij} = vs_{ij} (1 + \varphi(vp_{11}, \dots, vp_{ij}, \dots, vp_{mq})). \quad (9)$$

Во втором случае коррекция веса общего показателя вычисляется в пределах i -го объекта по q системам, т. е. в модели для каждой строки таблицы 1 учитываются веса $vp_{i1}, vp_{i2}, \dots, vp_{iq}$:

$$vs'_{ij} = vs_{ij} (1 + \varphi(vp_{i1}, vp_{i2}, \dots, vp_{iq})). \quad (10)$$

По второму варианту коррекция веса общего ij -го показателя i -го объекта осуществляется без учета влияния систем остальных объектов, что позволяет автономно изменять веса общих показателей по каждому объекту сравнения.

Далее в работе для коррекции весов общих показателей использованы локальная модель (вычисления по (8)) и глобальная модель второго типа (вычисления по (10)). На основе математических моделей (8) и (10) формируется таблица скорректированных весов общих показателей (табл. 2).

В качестве функции свертки в (10), в общем случае, целесообразно использовать средние показатели: среднеарифметическое значение (MA), среднегеометрическое значение (MGM), среднегармоническое значение (MHR) и др.

Таблица 2 – Матрица скорректированных весов общих показателей

Objects	1		2		...	q	
	s	p ₁	s	p ₂		s	p _m
O ₁	vs ₁₁		vs ₁₂		...	vs _{1q}	
O ₂	vs ₂₁		vs ₂₂		...	vs _{2q}	
O ₃	vs ₃₁		vs ₃₂		...	vs _{3q}	
...					...		
O _m	vs _{m1}		vs _{m2}		...	vs _{mq}	

Согласно [27] для различных правил вычисления среднего справедливо правило мажорантности:

$$MA \leq MGM \leq MHR.$$

Табл. 3 содержит типовые формулы вычисления средних весов показателей.

На основе таблицы скорректированных весов (табл. 3) вычисляется вектор поправочных коэффициентов для объектов O₁ – O_m (как MA/MGM и др.) – VCC_{MA} = (vcc_{ma1} ... vcc_{ma_m}) или

$$VCC_{MGM} = (vcc_{mgm1} \dots vcc_{mgm_m}).$$

Для повышения обоснованности принимаемых решений полученный вектор VCC_{MA} или VCC_{MGM} вводится на этапе составления сводной таблицы (предпоследний этап классического МАИ).

Таблица 3 – Основные формулы средних весов показателей

Name	Formula
MA	$y = \frac{1}{mq} \sum_{i=1}^m \sum_{j=1}^q v p_{ij}^*$
MGM	$y = mq \sqrt[m]{\prod_{i=1}^m \prod_{j=1}^q v p_{ij}^*}$
MHR	$y = mq \frac{1}{\sum_{i=1}^m \sum_{j=1}^q \frac{1}{v p_{ij}^*}}$

Для этого на основе весов критериев (w₁, w₂, ... w_q)^T, весов сравниваемых объектов O_{ij} (i=1–m, j=1–q) и вектора VCC_{MA} или VCC_{MGM} вычисляется итоговый вес объекта, определяется по одной из формул:

$$V_j = \sum_{i=1}^{i=q} O_{ij} \cdot w_i \cdot vcc_{mgm_j}, \quad (11)$$

$$V_j = \sum_{i=1}^{i=q} O_{ij} \cdot w_i \cdot vcc_{ma_j}. \quad (12)$$

Таким образом, итоговые веса сравниваемых инновационных объектов по критериям K₁–K_z вычисляются как произведение не двух, а трех множителей, что определяет модификацию МАИ на этапе вычисления итоговых весов инновационных объектов с учетом их индивидуальных особенностей.

4 РЕЗУЛЬТАТЫ

Как показано выше, сравниваемые инновационные объекты (на примере БЛА) имеют следующие количественные характеристики [23, 26]: m=3; q=5; s=6; d=3.

Пусть значения общих и индивидуальных показателей (s, p) объектов O₁, O₂, O₃ представлены в табл. 4.

Согласно табл. 4 имеется S=66 общих показателей и P=37 индивидуальных показателей. При этом 21 индивидуальный показатель (57%) описывает объект O₃, что определяет его как высоко инновационный среди сравниваемых объектов.

Таблица 4 – Матрица показателей систем для сравниваемых объектов

Object	T		C		L		E		N		Всего
	s	p ₁	s	p ₂	s	p ₃	s	p ₄	s	p ₅	
O ₁	4	1	3	1	6	2	5	2	4	1	7
O ₂	4	2	4	2	3	2	6	1	4	2	9
O ₃	5	5	5	3	4	5	4	4	5	4	21
Σ		8		6		9		7		7	37

В табл. 5 на основе (6) и (7) получена матрица весов общих и индивидуальных показателей объектов. В табл. 6 на основе (8) получена матрица скорректированных весов общих показателей.

Таблица 5 – Матрица весов показателей для сравниваемых объектов

Objects	Веса системы T		Веса системы C		Веса системы L		Веса системы E		Веса системы N	
	s	p ₁	s	p ₂	s	p ₃	s	p ₄	s	p ₅
O ₁	0,677	0,107	0,500	0,08	1,0	0,240	0,833	0,187	0,667	0,093
O ₂	0,667	0,213	0,667	0,16	0,500	0,240	1,000	0,093	0,667	0,187
O ₃	0,883	0,533	0,883	0,240	0,667	0,600	0,667	0,373	0,833	0,373

Таблица 6 – Матрица скорректированных весов общих показателей на основе локальной модели

Objects	Весы системы T		Весы системы C		Весы системы L		Весы системы E		Весы системы N	
	s	p ₁	s	p ₂	s	p ₃	s	p ₄	s	p ₅
O ₁	0,738		0,540		1,240		0,989		0,729	
O ₂	0,809		0,773		0,620		1,093		0,791	
O ₃	1,278		1,033		1,067		0,916		1,144	

На основе индивидуальных весов систем T, C, L, E, N из табл. 5 по формулам вычисления средних, приведенных в табл. 3, рассчитаны средние значения весов индивидуальных показателей (табл. 7). Эти значения необходимы для последующего вычисления вектора поправочных коэффициентов.

Таблица 7 – Средние веса индивидуальных показателей

Object	Весы индивидуальных показателей					V _{ma}	V _{mgm}	V _{mgr}
	vp1	vp2	vp3	vp4	vp5			
O ₁	0,107	0,08	0,240	0,187	0,093	0,141	0,129	0,119
O ₂	0,213	0,16	0,240	0,093	0,187	0,179	0,170	0,160
O ₃	0,533	0,240	0,600	0,373	0,373	0,424	0,403	0,382

Согласно правила мажорантности из трех средних для глобальной модели коррекции весов выбирается VCC_{MGM} для расчетов по (10) и вычисляются скорректированные веса общих показателей (табл. 8).

По матрицам скорректированных весов общих показателей на основе локальной и глобальной моделей (табл. 6 и табл. 8) как средние арифметическое и геометрическое по каждому объекту O₁–O₃ вычисляются VCC_{MA} и VCC_{MGM} соответственно (табл. 9).

Данные вектора представляют собой формализованную дополнительную информацию, которая далее будет введена в сводную таблицу (предпоследний этап МАИ) при расчете итоговых весов по выражениям (11) или (12).

Пусть для МАИ рассматриваемые технические объекты O₁–O₃ оцениваются по следующим критериям: K₁, K₂, K₃, K₄, K₅ [23–26].

Для данных объектов задана матрица парных сравнений критериев A(5×5) в следующем виде [26] (табл. 10).

Таблица 8 – Матрица скорректированных весов общих показателей на основе глобальной модели

Objects	Весы системы T		Весы системы C		Весы системы L		Весы системы E		Весы системы N	
	s	p ₁	s	p ₂	s	p ₃	s	p ₄	s	p ₅
O ₁	0,753		0,565		1,129		0,940		0,753	
O ₂	0,780		0,780		0,585		1,170		0,780	
O ₃	1,239		1,239		0,936		0,936		1,169	

Таблица 9 – Вектора поправочных коэффициентов

Objects	VCC _{MA}	VCC _{MGM}
O ₁	0,813	0,806
O ₂	0,804	0,799
O ₃	1,081	1,095

Вектор весов критериев $\bar{w} = (\bar{w}_1, \bar{w}_2, \dots, \bar{w}_m)^T$ рассматривается как MGM сравниваемых критериев.

Таблица 10 – Матрица парных сравнений критериев

Критерии	K ₁	K ₂	K ₃	K ₄	K ₅	Вектор весов критериев w
K ₁	1	2	3	5	6	0,44
K ₂	0,50	1	2	3	4	0,26
K ₃	0,33	0,5	1	2	3	0,16
K ₄	0,2	0,33	0,5	1	2	0,09
K ₅	0,16	0,25	0,33	0,5	1	0,06

Проверка индекса согласованности для табл. 10 показала, что исходную матрицу парных сравнений критериев A можно считать согласованной и использовать в дальнейших расчетах [13, 26].

Пусть получены по каждому критерию K_x (x=1–5) значения весов объектов-альтернатив (табл. 11). На основе данных из табл. 10 и табл. 11 вычислены итоговые веса объектов (табл. 12) как сумма произведений двух множителей: веса критерия K_x (x=1–5) и веса объекта O₁–O₃.

Таблица 11 – Весы объектов по критериям сравнения

Objects	Критерии				
	K ₁	K ₂	K ₃	K ₄	K ₅
O ₁	0,74	0,72	0,64	0,07	0,74
O ₂	0,17	0,17	0,26	0,28	0,17
O ₃	0,09	0,11	0,10	0,65	0,09

Таблица 12 – Итоговые веса объектов

Objects	Критерии					Итоговый вес
	K ₁	K ₂	K ₃	K ₄	K ₅	
O ₁	0,32	0,18	0,10	0,01	0,04	0,66
O ₂	0,07	0,04	0,04	0,03	0,01	0,19
O ₃	0,04	0,03	0,02	0,06	0,005	0,15

Таким образом, рассчитанная по стандартному МАИ упорядоченная последовательность летательных аппаратов имеет вид $O_1 \rightarrow O_2 \rightarrow O_3$.

В данных расчетах (табл. 12) не учитываются индивидуальные особенности инновационных объектов, т.е. вектор поправочных коэффициентов из табл. 9.

Таблица 13 является расширением стандартного МАИ. Она содержит сводную матрицу расчета итоговых весов объектов с учетом вектора поправочных коэффициентов на основе МА.

Таблица 13 – Сводная матрица весов объектов с учетом вектора поправочных коэффициентов

Objects	K_1	K_2	K_3	K_4	K_5	VCC_{MA}	Итоговый вес
	0,44	0,26	0,16	0,09	0,06		
O_1	0,74	0,72	0,64	0,07	0,74	0,813	0,55
O_2	0,17	0,17	0,26	0,28	0,17	0,804	0,15
O_3	0,09	0,11	0,10	0,65	0,09	1,081	0,16

По модифицированному МАИ получена измененная упорядоченная последовательность $O_1 \rightarrow O_3 \rightarrow O_2$. Она более точно отражает места в ранжировании объектов-альтернатив, так как учитывает индивидуальные показатели сравниваемых объектов.

Таким образом, объект, имеющий большее количество индивидуальных показателей, имеет более высокий итоговый вес.

Для проверки влияния количества индивидуальных показателей объектов было выполнено моделирование расчета выбранного VCC_{MA} в зависимости от количества индивидуальных показателей для фиксированного количества общих показателей $S=66$.

Пропорция распределения индивидуальных показателей по системам N_T, N_C, N_L, N_E, N_N имеет вид:

$$\begin{aligned}
 N_T &= \{(1,2,2), (1,1,3), (1,2,3), (1,2,5), (2,2,5), (3,2,5)\}; \\
 N_C &= \{(0,0,3), (0,1,3), (1,1,3), (1,2,3), (2,3,5), (2,3,5)\}; \\
 N_L &= \{(1,0,2), (1,1,4), (1,2,4), (2,2,5), (3,2,5), (3,2,5)\}; \\
 N_E &= \{(0,1,2), (1,1,3), (2,1,3), (2,1,4), (2,2,5), (2,2,5)\}; \\
 N_N &= \{(0,1,3), (0,2,3), (1,2,3), (1,2,4), (2,3,4), (2,3,5)\}.
 \end{aligned}$$

На основе общих показателей, имеющих распределение согласно табл. 4, и матрицы значений по системам N_T, N_C, N_L, N_E, N_N было сформировано 6 таблиц, по которым выполнены расчеты векторов поправочных коэффициентов. В частности, четвертый столбец индивидуальных показателей по системам N_T, N_C, N_L, N_E, N_N обеспечивает формирование табл. 4.

В табл. 14 отражена доля индивидуальных показателей к суммарному количеству общих показателей ($S=66$).

Табл. 15 содержит значения векторов поправочных коэффициентов для шести матриц, причем четвертый столбец табл. 15 соответствует расчетному значению из табл. 9.

Таблица 14 – Доля индивидуальных показателей для шести матриц исходных данных

Показатели	Номер матрицы					
	1	2	3	4	5	6
$\sum_{i=1}^q \sum_{j=1}^m P_{ij}^*$	18	25	30	37	47	50
$\sum_{i=1}^q \sum_{j=1}^m P_{ij}^*/S$ (в %)	27,2	37,8	45,4	56	71,2	75,7

Таблица 15 – Таблица значений вектора поправочных коэффициентов VCC_{MA}

Objects	Матрицы моделируемых данных					
	1	2	3	4	5	6
O_1	0,729	0,744	0,782	0,813	0,870	0,885
O_2	0,713	0,737	0,771	0,804	0,841	0,856
O_3	0,849	0,926	0,958	1,081	1,122	1,137

Расчеты итоговых весов по модифицированному МАИ для шести матриц приведены в табл. 16, причем четвертый столбец табл. 16 соответствует расчетному значению из табл. 13.

Таблица 16 – Таблица итоговых весов сравниваемых объектов с учетом VCC_{MA}

Objects	Матрицы данных					
	1	2	3	4	5	6
O_1	0,479	0,489	0,514	0,545	0,572	0,582
O_2	0,137	0,142	0,148	0,154	0,162	0,165
O_3	0,127	0,139	0,144	0,162	0,168	0,171

Пусть по критериям K_1, K_2, K_3, K_4, K_5 [23–26] сравниваются новые пять инновационных объектов $ИО_1$ – $ИО_5$ и пусть для них матрица парных сравнений критериев имеет вид согласно табл. 10.

В рамках классического МАИ пусть по каждому критерию K_x ($x=1$ – 5) получены значения весов инновационных объектов (табл. 17). На основе данных из табл. 10 и табл. 17 вычислены итоговые веса инновационных объектов (табл. 18).

Таблица 17 – Веса ИО по критериям сравнения

Objects	Критерии				
	K_1	K_2	K_3	K_4	K_5
	0,44	0,26	0,16	0,09	0,06
$ИО_1$	0,45	0,16	0,26	0,25	0,58
$ИО_2$	0,25	0,52	0,28	0,41	0,20
$ИО_3$	0,31	0,33	0,47	0,34	0,22
$ИО_4$	0,27	0,43	0,28	0,21	0,38
$ИО_5$	0,10	0,59	0,48	0,26	0,43

Таблица 18 – Сводная матрица расчета весов ИО

Objects	K_1	K_2	K_3	K_4	K_5	Итог. вес
$ИО_1$	0,196	0,040	0,040	0,023	0,033	0,332
$ИО_2$	0,108	0,132	0,043	0,037	0,011	0,331
$ИО_3$	0,136	0,084	0,073	0,031	0,013	0,336
$ИО_4$	0,118	0,110	0,043	0,019	0,022	0,320
$ИО_5$	0,043	0,152	0,075	0,024	0,025	0,318

Анализ итоговых весов $ИО_1-ИО_5$, полученных по классическому МАИ, показывает, что они являются практически неразличимыми, что затрудняет выбор лучшей альтернативы.

Предлагаемый модифицированный МАИ на основе структурно-функциональной декомпозиции инновационных объектов на q систем позволяет ввести дополнительную информацию об $ИО_1-ИО_5$ в виде матрицы общих и индивидуальных показателей систем.

Пусть для $ИО_1-ИО_5$ значения общих и индивидуальных показателей (s, p) представлены в табл. 19. При этом принимается $m=5; q=5; s=6; d=6$.

Согласно табл. 19 имеется $S=105$ общих показателей и $P=85$ индивидуальных показателей. Эти значения подчеркивают статус сравниваемых объектов как инновационных, т.е. объектов, имеющих значительную долю индивидуальных показателей в их описании.

Таблица 19 – Матрица показателей систем для сравниваемых объектов

Object	T		C		L		E		N		Σ
	s	p_1	s	p_2	s	p_3	s	p_4	s	p_5	
$ИО_1$	4	3	3	4	6	4	5	2	4	3	16
$ИО_2$	5	4	5	3	4	3	3	4	5	2	16
$ИО_3$	5	4	5	4	4	4	4	3	5	2	17
$ИО_4$	4	4	4	5	3	3	3	3	4	3	18
$ИО_5$	4	3	5	4	3	3	4	4	4	4	18
Всего											85

На основе выражений (6), (7) составляется матрица весов общих и индивидуальных показателей, а по ней на основе локальной модели коррекции весов – матрица скорректированных весов (табл.20). Последняя матрица является основой для вычисления вектора поправочных коэффициентов на основе среднегеометрического значения (табл. 21).

Таблица 22 содержит сводную матрицу расчета итоговых весов инновационных объектов как сумму произведений трех множителей: веса критерия ($K=1-5$), веса объекта $ИО_1-ИО_5$ и элемента вектора поправочных коэффициентов VCC_{MGM} .

Модифицированный МАИ позволил ранжировать инновационные объекты $ИО_1-ИО_5$ и выделить лучшую альтернативу – инновационный объект 3 как имеющий больший итоговый вес.

Таблица 20 – Матрица скорректированных весов показателей для $ИО_1-ИО_5$ на основе локальной модели

Objects	Весы системы T		Весы системы C		Весы системы L		Весы системы E		Весы системы N	
	s	p	s	p	s	p	s	p	s	p
$ИО_1$	0,87		0,72		1,38		0,98		0,82	
$ИО_2$	1,17		1,11		0,86		0,68		0,96	
$ИО_3$	1,17		1,20		0,92		0,84		0,96	
$ИО_4$	0,93		1,04		0,64		0,63		0,82	
$ИО_5$	0,87		1,20		0,64		0,90		0,87	

Таблица 21 – Вектор поправочных коэффициентов

Objects	VCC_{MGM}
$ИО_1$	0,930
$ИО_2$	0,937
$ИО_3$	1,010
$ИО_4$	0,798
$ИО_5$	0,880

Таблица 22 – Сводная матрица весов инновационных объектов с учетом вектора поправочных коэффициентов

Objects	K_1	K_2	K_3	K_4	K_5	VCC_{MGM}	Итоговый вес
	0,44	0,26	0,16	0,09	0,06		
$ИО_1$	0,45	0,16	0,26	0,25	0,58	0,930	0,30
$ИО_2$	0,25	0,52	0,28	0,41	0,20	0,937	0,31
$ИО_3$	0,31	0,33	0,47	0,34	0,22	1,010	0,34
$ИО_4$	0,27	0,43	0,28	0,21	0,38	0,798	0,25
$ИО_5$	0,10	0,59	0,48	0,26	0,43	0,880	0,28

5 ОБСУЖДЕНИЕ

В работе были рассмотрены технические инновационные объекты, описываемые общими и индивидуальными показателями с пропорцией 1:2 с количеством критериев сравнения, равным 5, что соответствует классу объектов средней организационной сложности. Исходные значения общих и индивидуальных показателей распределены таким образом, что содержат представителя типового объектов (количество индивидуальных показателей существенно меньше 50%), представителя инновационного объекта (количество индивидуальных показателей больше 50%), представителя объекта в пограничной зоне (количество индивидуальных показателей около 50%).

Классический МАИ определяет объект O_3 как наименее приоритетный среди сравниваемых объектов O_1-O_3 , так как не учитывает индивидуальные особенности инновационного объекта. Напротив, модифицированный МАИ позволяет устранить это ограничение и ввести в расчеты дополнительную информацию. Расширение вычислений в сводной таблице итоговых весов показало, что наименее приоритетным стал объект O_2 , так как он имеет одновременно пониженные значения в матрице парных сравнений по критериям и низкие значения индивидуальных показателей.

Сравнение пяти инновационных объектов $ИО_1$ – $ИО_5$ классическим и разработанным модифицированным МАИ показало, что классический МАИ не может ранжировать и выделить лучшую альтернативу в силу нивелирования весов инновационных объектов по критериям при их свертке в итоговый вес. Разработанный модифицированный МАИ за счет учета индивидуальных показателей инновационных объектов позволил разнести их на оси измерения итогового веса на расстояния, достаточные для ранжирования и выделения лучшей альтернативы.

В целом, модифицированный метод анализа иерархий отличаются:

- применение структурно-функциональной декомпозиции объектов на системы;
- описание систем объектов матрицей общих и индивидуальных показателей;
- низкая вычислительная сложность моделей коррекции весов и вычисления вектора поправочных коэффициентов.

Данные особенности позволяют применять модифицированный МАИ для сравнения инновационных объектов из различных предметных областей (подвижная техника, космические аппараты, биотехнические изделия, экономические системы и др. [28–34]).

ВЫВОДЫ

В работе решена актуальная задача разработки модифицированного МАИ для ранжирования инновационных объектов с применением моделей коррекции весов.

Научная новизна полученных результатов заключается в том, что предложен модифицированный МАИ с учетом структурно-функциональной организации инновационных объектов. На основе данного метода, расширенного моделями коррекции весов и вектором поправочных коэффициентов, получена последовательность ранжирования объектов, учитывающая влияние индивидуальных показателей на выбор лучшего объекта. Модифицированный метод объективно отражает измененный порядок следования объектов-альтернатив с учетом баланса парных оценок альтернатив и их структурно-функционального описания в виде матрицы общих и индивидуальных показателей.

Практическая значимость полученных результатов заключается в том, что согласно проведенного моделирования ценным для практики принятия решений следует считать долю индивидуальных показателей в их общем количестве более 55%. В этом случае объекты предпочтительнее оценивать на основе модифицированного МАИ, т.е. с использованием вектора поправочных коэффициентов. Если доля количества индивидуальных показателей к количеству общих показателей меньше 45%, то предпочтительным является классический МАИ. Влияние индивидуальных показателей незначительное. Диапазон [45%–55%] является пограничной зоной для принятия решений по инновационным объектам на основе поправочных

коэффициентов. В целом, модели коррекции весов, введенные в состав МАИ, позволяют получить уточненный порядок следования инновационных объектов с учетом их структурно-функциональной декомпозиции.

Полученный вектор поправочных коэффициентов может иметь самостоятельное применение и использоваться при решении различных расчетно-аналитических задач и прикладных задач принятия решений.

Перспективы дальнейших исследований заключаются в применении разработанных моделей коррекции весов и модифицированного метода для сравнения различных инновационных объектов – реконфигурируемые подвижные роботы, биотехнические изделия, мехатронные приводы и др.

БЛАГОДАРНОСТИ

Автор выражает благодарность Емельянову Сергею Геннадьевичу, ректору ЮЗГУ, доктору технических наук, профессору, члену-корреспонденту РААСН за техническую помощь в предоставлении научно исследовательской и лабораторной базы для проведения исследований в данной области.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Юдин Д. Б. Математические методы управления в условиях неполной информации / Д. Б. Юдин. – М. : URSS, 2010. – 400 с.
2. Зак Ю. А. Принятие многокритериальных решений / Ю. А. Зак. – М. : Экономика, 2011. – 235 с.
3. Кук Ю. В. Спектральный метод распознавания состояний динамических систем / Ю. В. Кук, Е. И. Лаврикова // Комп'ютерні засоби, мережі та системи. – 2007. – № 6. – С. 133–140.
4. Saaty T. L. Decision making with the Analytic Hierarchy Process / T. L. Saaty // International Journal of Services Sciences. – 2008. – Vol. 1, No. 1. – P. 83–98.
5. Дубровин В. И. Метод получения вектора приоритетов из нечетких матриц парных сравнений / В. И. Дубровин, Н. А. Миронова // Искусственный интеллект. – 2009. – № 3. – С. 464–470.
6. Щеглов С. Н. Анализ моделей принятия решений в новых информационных технологиях в условиях нечеткости и неопределенности / С. Н. Щеглов // Известия ЮФУ. Технические науки. – 2012. – Т. 132, № 7. – С. 126–131.
7. Groselj P. Evaluation of several approaches for deriving weights in fuzzy group analytic hierarchy process / P. Groselj, Stirn L. Zadnik // Journal of Decision Systems. – 2018. – Vol. 27. – P. 217–226.
8. Sieni E. Migration NSGA : Method to Improve a Non-Elitist Searching of Pareto front with Application in Magnetics / E. Sieni, M. Forzan, P. Di Barba // Inverse Problems in Science and Engineering. – 2016. – Vol. 24, No. 1. – P. 543–566.
9. Wu S. A consistency improving method in the analytic hierarchy process based on directed circuit analysis / S. Wu, X. Liu, Z. Li, Y. Zhou // Journal of Systems En-

- gineering and Electronics. – 2019. – Vol. 30, No. 6. – P. 1160–1181.
10. Подиновский В. В. Парето-оптимальные решения многокритериальных задач / В. В. Подиновский, В. Д. Ногин. – М. : ФИЗМАТЛИТ, 2007. – 256 с.
 11. Машуров А. Ю. Метод анализа иерархий для определения лучшей альтернативы / А. Ю. Машуров, А. Ф. Винокуров, А. И. Левочки // Молодой ученый. – 2019. – № 17. – С. 34–38.
 12. Watkins D. S. Fundamentals of Matrix Computations / D. S. Watkins // Wiley-Interscience. – 2002. – P. 633.
 13. Саати Т. Л. Принятие решений. Метод анализа иерархий / Т. Л. Саати. – М. : Радио и связь, 1993. – 320 с.
 14. Саати Т. Л. Принятие решений при зависимостях и обратных связях : аналитические сети / Т. Л. Саати. – М. : Книжный дом «ЛИБРОКОМ», 2009. – 360 с.
 15. Comparison of the multicriteria decision-making methods for equity portfolio selection: The U.S. evidence / [E. Patari, V. Karell, P. Luukka et al] // European Journal of Operational Research, 2018. – P. 655–672.
 16. Миронова Н. А. Интеграция модификаций метода анализа иерархий для систем поддержки принятия групповых решений / Н. А. Миронова // Радіоелектроніка, інформатика, управління. – 2011. – № 2. – С. 47–57.
 17. Миронова Н. А. Архитектура групповой системы поддержки принятия решений с возможностью синтеза метода принятия групповых решений / Н. А. Миронова, А. А. Скрипник // Системи обробки інформації. – 2012. – Т. 8. – С. 33.
 18. Черкашин А. К. Математические аспекты реализации метода анализа иерархий / А. К. Черкашин // Информационные и математические технологии в науке и управлении. – 2020. – № 1 (17). – С. 5–24.
 19. Подиновский В. В. О некорректности метода анализа иерархий / В. В. Подиновский, О. В. Подиновская // Проблемы управления. – 2011. – № 1. – С. 8–13.
 20. Миронова Н. А. Экспертная система выбора модификации метода анализа иерархий / Н. А. Миронова // Автоматизированные системы управления и приборы автоматики. – 2011. – С. 62–70.
 21. Mousavia S. J. Inferring operating rules for reservoir operations using fuzzy regression and ANFIS / S. J. Mousavia, K. Ponnambalamb, F. Karray // Fuzzy Sets and Systems. – 2007. Vol. 158. – P. 1064–1082.
 22. Тутьгин А. Г. Преимущества и недостатки метода анализа иерархий / А. Г. Тутьгин, В. Б. Коробов // Известия Российского государственного педагогического университета им. А. И. Герцена. Естественные и точные науки. – 2010. – № 122. – С. 108–115.
 23. Пшихопов В. Х. Оценивание и управление в сложных динамических системах / В. Х. Пшихопов, М. Ю. Медведев. – М. : Физматлит, 2009. – 294 с.
 24. Methods of Recognition Based on the Function of Rival Similarity / [N. G. Zagoruiko, I. A. Borisova, V. V. Dyubanov et al] // Pattern Recognition and Image Analysis, 2008. – Vol. 18, No. 1. – P. 1–6.
 25. Тетерин Д. П. Синтез требований к бортовому информационно-измерительному и моделирующему комплексу / Д. П. Тетерин // Информационно-управляющие системы. – 2009. – № 1. – С. 10–14.
 26. Локальная математическая модель подготовки данных для сравнения инновационных объектов / [Е. А. Титенко, О. Г. Добросердов, Л. А. Лисицин и др.] // Инновации. – 2019. – № 12. – С. 74–81.
 27. Анфилатов В. С. Системный анализ в управлении: учебное пособие / В. С. Анфилатов, А. А. Емельянов, А. А. Кукушкин. – М. : Финансы и статистика, 2002. – 368 с.
 28. Ostroukh A.V., Elhadi H. Comparative study of routing protocols in vehicular ad-hoc networks (vanets) / A. V. Ostroukh, H. Elhadi // International Journal of Advanced Studies (iJAS). – 2014. – Vol. 4, No 2. – P. 9–14.
 29. Трахтенгерц Э. А. Компьютерные системы и методы поддержки информационного управления / Э. А. Трахтенгерц. – М. : СИНТЕГ, 2010. – 135 с.
 30. Титенко Е. А. Модифицированная производственная система и специализированное производственное устройство для поддержки решения проблемно-поисковых задач / Е. А. Титенко, Т. А. Мирталибов // Вісник ДУІКТ. – 2013. – № 3. – С. 13–19.
 31. Математические модели коррекции весов инновационных объектов / [Е. А. Титенко, В. В. Варганов, С. Н. Фролов и др.] // Робототехника и искусственный интеллект : XI Всероссийская научно-техническая конференция с международным участием: тезисы докладов. – Железнодорожск : Сибирский федеральный университет, 2019. – С. 154–160.
 32. Остроух А. В. Интеллектуальные информационные системы и технологии : монография / А. В. Остроух, Н. Е. Суркова. – Красноярск : Научно-инновационный центр, 2015. – 370 с.
 33. Andronov V. G. Autonomous navigation and attitude control of spacecrafts on near-earth circular orbits. / V. G. Andronov, S. G. Emelyanov // Journal of Applied Engineering Science. – 2018. – Vol. 16, № 1. – P. 107–110.
 34. Многоагентные системы в технической диагностике сложных технических объектов / [Т. А. Мирталибов, С. Н. Фролов, А. Л. Ханис и др.] // Известия Юго-Западного университета. – 2015. – № 3. – С. 18–25.

Статья поступила в редакцию 23.03.2020.
После доработки 21.09.2020.

УДК 004.023

МОДЕЛІ ОБЧИСЛЕННЯ ВАГ ДЛЯ ОЦІНКИ ІННОВАЦІЙНИХ ТЕХНІЧНИХ ОБ'ЄКТІВ

Титенко Є. А. – канд. техн. наук, доцент, провідний науковий співробітник центру перспективних досліджень і розробок, південно-західний державний університет, Курськ, Росія.

Фролов С. Н. – канд. техн. наук, провідний науковий співробітник центру перспективних досліджень і розробок, південно-західний державний університет, Курськ, Росія.

Ханіс А. Л. – канд. військ. наук, доцент кафедри інформаційної безпеки, південно-західний державний університет, Курськ, Росія.

Добросердов О. Г. – д-р техн. наук, старший науковий співробітник, радник при ректораті, південно-західний державний університет, Курськ, Росія.

Захаренков А. І. – д-р техн. наук, професор, перший заступник генерального директора, АТ «Воентелеком», Москва, Росія.

Попов О. М. – генеральний директор, АТ «Авіаавтоматика» ім. В. В. Тарасова, Курськ, Росія.

Дронова Т. А. – д-р мед. наук, професор кафедри пропедевтики внутрішніх хвороб, Курський державний медичний університет, Курськ, Росія.

АНОТАЦІЯ

Актуальність. Актуальність роботи пов'язана з багатокритеріальним порівнянням інноваційних об'єктів, під якими розуміються об'єкти з частково співпадаючими індивідуальними показниками, що мають невизначені, суб'єктивні значення і критерії їх оцінки. Порівнювані інноваційні об'єкти описуються загальними та індивідуальними показниками, причому кількість індивідуальних показників є значною (більше 50%). Метод аналізу ієрархій є найбільш підходящим математичним апаратом для порівняння інноваційних об'єктів, так як він поєднує наочність опису предметної області, чисельну результативність, природний облік варіативності вихідних парних оцінок. Проте, даний метод слабо враховує індивідуальні показники інноваційних об'єктів при складанні матриць парних порівнянь, що не дозволяє отримати підсумкові ваги, відповідні індивідуальним особливостям інноваційних об'єктів.

Ціль. Розробка модифікованого методу аналізу ієрархій, що забезпечує ранжування інноваційних об'єктів з частково співпадаючими індивідуальними показниками.

Метод. В роботі розроблено модифікований метод аналізу ієрархій, що включає типову послідовність етапів від побудови ієрархії предметної області і формування критеріїв оцінки інноваційних об'єктів до етапу обчислення підсумкових ваг інноваційних об'єктів, що відрізняється по-перше, введенням на етапі побудови ієрархії предметної області матриці ваг загальних та індивідуальних показників по системам об'єктів, по-друге розрахунком поправочних коефіцієнтів на основі локальної або глобальної моделей корекції ваг показників систем об'єктів на етапі складання зведеної таблиці ваг, по-третє обчисленням підсумкових ваг інноваційних об'єктів з урахуванням поправочних коефіцієнтів, що дозволило врахувати індивідуальні особливості порівнюваних інноваційних об'єктів.

Результати. У роботі розглянуто інноваційні об'єкти із загальними та індивідуальними показниками в пропорції 1: 2 з кількістю критеріїв порівняння, що дорівнює 5, що відповідає класу об'єктів середньої організаційної складності. Порівнювані об'єкти містять Типовий об'єкт (кількість індивідуальних показників істотно менше 50%), явно інноваційний об'єкт (кількість індивідуальних показників більше 50%), об'єкт в прикордонній зоні (кількість індивідуальних показників близько 50%). Класичний метод не чутливий до індивідуальних особливостей інноваційних об'єктів, що визначило мінімальну вагу для явно інноваційного об'єкта. Модифікований метод, навпаки, визначив мінімальну вагу для об'єкта в прикордонній зоні, так як він має одночасно знижені значення в матриці парних порівнянь за критеріями і низькі значення індивідуальних показників.

Висновки. Розроблений модифікований метод аналізу ієрархій і моделі корекції в його складі об'єктивно відображають порядок ранжирування об'єктів з урахуванням їх опису у вигляді матриці загальних і індивідуальних показників. Згідно з моделюванням, цінним для практики прийняття рішень слід вважати частку індивідуальних показників в їх загальній кількості більше 55%. У цьому випадку об'єкти краще оцінювати на основі модифікованого методу аналізу ієрархій. Отриманий на основі моделі корекції ваг вектор поправочних коефіцієнтів має самостійне значення при вирішенні різних розрахунково-аналітичних задач і прикладних задач прийняття рішень.

КЛЮЧОВІ СЛОВА: інноваційний високотехнологічний об'єкт, модель корекції ваг, Розширена зведена таблиця, вектор поправочних коефіцієнтів, метод аналізу ієрархій.

УДК 004.023

MODELS FOR CALCULATING WEIGHTS FOR ESTIMATION INNOVATIVE TECHNICAL OBJECTS

Titenko E. A. – PhD, Associate Professor, Leading Researcher of the Center for Advanced Research and development, South-Western state University, Kursk, Russia.

Frolovs N. S. – PhD, Leading Researcher at the Center for Advanced Research and Development, South-Western state University, Kursk, Russia.

Khanis A. L. – PhD, Associate Professor of the Department of Information security, South-Western state University, Kursk, Russia.

Dobroserdov O. G. – Dr. Sc., Senior Researcher, Adviser to the rector, South-Western state University, Kursk, Russia.

Zakharenkov A. I. – Dr. Sc., Professor, first Deputy General Director, Voentelcom JSC, Moscow, Russia.

Popov A. N. – General Director, V. V. Tarasov Aviaavtomatika JSC, Kursk, Russia.

Dronova T. A. – Dr. Sc., Professor of Propaedeutics of Internal Diseases, Kursk State Medical University of the Ministry of health of the Russian Federation, Kursk, Russia.

ABSTRACT

Relevance. The relevance of the work is associated with a multi-criteria comparison of innovative objects, which are understood as objects with partially identical individual indicators that have uncertain, subjective values and criteria for their evaluation. Compared innovation objects are described by General and individual indicators, and the number of individual indicators is significant (more than 50%). The hierarchy analysis method is the most suitable mathematical tool for comparing innovative objects, since it combines visual description of the subject area, numerical performance, and natural accounting for the variability of the initial pair estimates. However, this method does not take into account the individual indicators of innovative objects when compiling matrices of paired comparisons, which does not allow you to get the final weights corresponding to the individual characteristics of innovative objects.

Object. Development of a modified hierarchy analysis method that provides ranking of innovative objects with partially identical individual indicators.

Method. In this work, a modified method of analysis of hierarchies, comprising a typical sequence of stages from building a hierarchy of the subject area and the establishment of criteria for evaluation of innovative objects to the stage computing the final weights of innovative objects, a first introduction on the stage of building a hierarchy of the weights matrix shared and individual performance in the system of objects, and secondly by calculating correction factors based on local or global models of correction of weights of parameters of systems of objects at the stage of preparation of the consolidated table of weights, thirdly the computation of the final weights of innovative features based on a correction factor that took into account individual features compare innovative.

Results. The paper considers innovative objects with General and individual indicators in the ratio of 1:2 with the number of comparison criteria equal to 5, which corresponds to the class of objects of average organizational complexity. The objects being compared contain a typical object (the number of individual indicators is significantly less than 50%), a clearly innovative object (the number of individual indicators is more than 50%), and an object in the border zone (the number of individual indicators is about 50%). The classical method is not sensitive to the individual characteristics of innovative objects, which determined the minimum weight for a clearly innovative object. The modified method, on the contrary, determined the minimum weight for the object in the border zone, since it has both reduced values in the matrix of paired comparisons by criteria and low values of individual indicators.

Conclusions. The developed modified method of hierarchy analysis and correction models in its composition objectively reflect the order of ranking of objects, taking into account their description in the form of a matrix of General and individual indicators. According to the modeling, the share of individual indicators in their total number of more than 55% should be considered valuable for decision-making practice. In this case, it is preferable to evaluate objects based on a modified hierarchy analysis method. The vector of correction coefficients obtained on the basis of the weight correction model has an independent value in solving various computational and analytical problems and applied decision-making problems.

KEYWORDS: innovative high-tech object, a model for weight correction, an extended summary table, a vector of correction coefficients, and a method for analyzing hierarchies.

REFERENCES

1. Judin D. B. *Matematicheskie metody upravlenija v uslovijah nepolnoj informacii*. Moscow, URSS, 2010, 400 p.
2. Zak Ju. A. *Prinjatje mnogokriterial'nyh reshnij*. Moscow, Jekonomika, 2011, 235 p.
3. Kuk Ju. V., Lavrikova E. I. *Spektral'nyj metod raspoznavanija sostojanij dinamicheskikh sistem, Komp'juterni zasobi, mrezi ta sistemi*, 2007, No. 6, pp. 133–140.
4. Saaty T. L. *Decision making with the Analytic Hierarchy Process*, *International Journal of Services Sciences*, 2008, Vol. 1, No. 1, pp. 83–98.
5. Dubrovin V. I., Mironova N. A. *Metod poluchenija vektora prioritetov iz nechetkih matric parnyh sravnenij, Iskustvennyj intellekt*, 2009, No. 3, pp. 464–470.
6. Shheglov S. N. *Analiz modelej prinjatija reshenij v novyh informacionnyh tehnologijah v uslovijah nechetkosti i neopredelennost, Izvestija JuFU. Tehniceskie nauk*, 2012. Vol. 132, No. 7, pp. 126–131.
7. Grosej P., Zadnik Stirn L. *Evaluation of several approaches for deriving weights in fuzzy group analytic hierarchy process*, *Journal of Decision Systems*, 2018, Vol. 27, pp. 217–226.
8. Sieni E., Forzan M., Barba P. *Di Migration NSGA : Method to Improve a Non-Elitist Searching of Pareto front with Application in Magnetics*, *Inverse Problems in Science and Engineering*, 2016, Vol. 24, No. 1, pp. 543–566.
9. Wu S., Liu X., Li Z., Zhou Y. *A consistency improving method in the analytic hierarchy process based on directed circuit analysis*, *Journal of Systems Engineering and Electronics*, 2019, Vol. 30, No. 6, pp. 1160–1181.
10. Podinovskij V. V., Nogin V. D. *Pareto-optimal'nye reshenija mnogokriterial'nyh zadach*. Moscow, FIZMATLIT, 2007, 256 p.
11. Mashurov A. Ju., Vinokurov A. F., Levochko A. I. *Metod analiza ierarhij dlja opredelenija luchshej al'ternativy, Molodoj uchenyj*, 2019, No. 17, pp. 34–38.
12. Watkins D. S. *Fundamentals of Matrix Computations*, *Wiley-Interscience*, 2002, P. 633.
13. Saati T. L. *Prinjatje reshenij. Metod analiza ierarhij*. Moscow, Radio i svjaz', 1993, 320 p.
14. Saati T. L. *Prinjatje reshenij pri zavisimostjah i obratnyh svjazjah : analiticeskie seti*. Moscow, Knizhnyj dom «LIBROKOM», 2009, 360 p.
15. Patari E., Karell V., Luukka P. et al *Comparison of the multicriteria decision-making methods for equity portfolio selection: The U.S. evidence*, *European Journal of Operational Research*, 2018, pp. 655–672.
16. Mironova N. A. *Integracija modifikacij metoda analiza ierarhii dlja sistem podderzhki prinjatija gruppyvnyh reshenij*, *Radio Electronics, Computer Science, Control*, 2011, No. 2, pp. 47–57.

17. Mironova N. A., Skripnik A. A. Arhitektura gruppovoj sistemy podderzhki prinjatija reshenij s vozmozhnost'ju sinteza metoda prinjatija gruppovyh reshenij, *Sistemi obrobki informacii*, 2012, Vol. 8, P. 33.
18. Cherkashin A. K. Matematicheskie aspekty realizacii metoda analiza ierarhij, *Informacionnye i matematicheskie tehnologii v nauke i upravlenii*, 2020, No. 1 (17), pp. 5–24.
19. Podinovskij V. V., Podinovskaja O. V. O nekorrektnosti metoda analiza ierarhij, *Problemy upravlenija*, 2011, No. 1, pp. 8–13.
20. Mironova N. A. Jekspertnaja sistema vybora modifikacii metoda analiza ierarhij, *Avtomatizirovannye sistemy upravlenija i pribory avtomatiki*, 2011, pp. 62–70.
21. Mousavia S. J., Ponnambalamb K., Karray F. Inferring operating rules for reservoir operations using fuzzy regression and ANFIS, *Fuzzy Sets and Systems*, 2007, Vol. 158, pp. 1064–1082.
22. Tutygin A. G., Korobov V. B. Preimushhestva i nedostatki metoda analiza ierarhij, *Izvestija Rossijskogo gosudarstvennogo pedagogicheskogo universiteta im. A. I. Gercena. Estestvennye i tochnye nauki*, 2010, No. 122, pp. 108–115.
23. Pshihopov V. H., Medvedev M. Ju. Ocenivanie i upravlenie v slozhnyh dinamicheskikh sistemah. Moscow, Fizmatlit, 2009, 294 p.
24. Zagoruiko N. G., Borisova I. A., Dyubanov V. V. et al. Methods of Recognition Based on the Function of Rival Similarity, *Pattern Recognition and Image Analysis*, 2008, Vol. 18, No. 1, pp. 1–6.
25. Teterin D. P. Sintez trebovanij k bortovomu informacionno-izmeritel'nomu i modelirujushhemu kompleksu, *Informacionno- upravljajushhie sistemy*, 2009, No. 1, pp. 10–14.
26. Titenko E. A., Dobroserdov O. G., Lisicin L. A. et al. Lokal'naja matematicheskaja model' podgotovki dannyh dlja sravnenija innovacionnyh ob'ektov, *Innovacii*, 2019, No. 12, pp. 74–81.
27. Anfilatov V. S. Emel'janov A. A., Kukushkin A. A. Sistemnyj analiz v upravlenii: uchebnoe posobie. Moscow, Finansy i statistika, 2002, 368 p.
28. Ostroukh A. V., Elhadi H. Comparative study of routing protocols in vehicular ad-hoc networks (vanets), *International Journal of Advanced Studies (iJAS)*, 2014, Vol. 4, No. 2, pp. 9–14.
29. Trahtengerc Je. A. Komp'juternye sistemy i metody podderzhki informacionnogo upravlenija. Moscow, SINTEG, 2010, 135 p.
30. Titenko E. A., Mirtalibov T. A. Modificirovannaja produkcionnaja sistema i specializirovannoe produkcionnoe ustrojstvo dlja podderzhki reshenija problemno-poiskovyh zadach, *Visnik DUIKT*, 2013, No. 3, pp. 13–19.
31. Titenko E. A., Varganov V. V., Frolov S. N. i dr. Matematicheskie modeli korrekcii vesov innovacionnyh ob'ektov, *Robototehnika i iskusstvennyj intellekt : XI Vserossijskaja nauchno-tehnicheskaja konferencija s mezhdunarodnym uchastiem: tezisy dokladov*. Zheleznogorsk, Sibirskij federal'nyj universitet, 2019, pp. 154–160.
32. Ostrouh A. V., Surkova N. E. Intellektual'nye informacionnye sistemy i tehnologii: Monografija. Krasnojarsk, Nauchno-innovacionnyj centr, 2015, 370 p.
33. Andronov V. G., Emelyanov S. G. Autonomous navigation and attitude control of spacecrafts on near-earth circular orbits, *Journal of Applied Engineering Science*, 2018, Vol. 16, No. 1, pp. 107–110.
34. Mirtalibov T. A., Frolov S. N., Hanis A. L. i dr. Mnogo-agentnye sistemy v tehnicheskoi diagnostike slozhnyh tehnicheskikh ob'ektov, *Izvestija Jugo-Zapadnogo universiteta*, 2015, No. 3, pp. 18–25.

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

CONTROL IN TECHNICAL SYSTEMS

УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

UDC 621.391:004.052

DEVELOPMENT AND RESEARCH OF A WIRELESS CONTROL SYSTEM FOR DEVICE “BIONIC EAR”

Arutyunov S. D. – Dr. Sc., Head of Department “Propaedeutics of dental diseases”, A. I. Yevdokimov Moscow State University of Medicine and Dentistry, Moscow, Russia.

Astashina N. B. – Dr. Sc., Head of Department “Orthopedic Dentistry”, E. A. Vagner Perm State Medical University, Perm, Russia.

Bajdarov A. A. – PhD, Vice Rector for Information Technology and Innovation Development, E. A. Vagner Perm State Medical University, Perm, Russia.

Gurko V. A. – Senior Lecturer of Department “Automatics and telemechanics”, Perm National Research Polytechnic University, Perm, Russia.

Mayorov P. V. – Student of Department “Automatics and telemechanics”, Perm National Research Polytechnic University, Perm, Russia.

Freyman V. I. – Dr. Sc., Professor of Department “Automatics and telemechanics”, Perm National Research Polytechnic University, Perm, Russia.

Yuzhakov A. A. – Dr. Sc., Head of Department “Automatics and telemechanics”, Perm National Research Polytechnic University, Perm, Russia.

ABSTRACT

Context. The analysis of reliability indicators of information transfer between elements of wireless control systems – a control device implemented on a smartphone and a communication processor of the developed hearing aid is carried out. The object of research is the reliability indicators of the wireless control channel. The subject of the study is a theoretical and experimental analysis of the dependencies of reliability indicators of the designed wireless control system.

Objective. The purpose of the work is to determine and study the analytical and experimental dependences of the reliability of transmission over the wireless control channel on the properties of the communication channel and settings of the system elements, forming recommendations for setting the parameters of the elements of the control system.

Methods. The elements of the reliability theory are used to determine the dependences of the bit and block error coefficients on the properties of the communication channel of the configuration of the elements of the control system. Analytical relationships are obtained to determine the reliability of transmission taking into account possible signal distortions. The dependences of reliability indicators on the parameters of the elements of the control system are investigated, illustrative examples are given. Simulated circuit models of a control system with selected Bluetooth wireless technology have been developed. Experimental studies have been carried out, on the basis of the data obtained, conclusions have been drawn and recommendations have been made on the choice of control system configurations in order to ensure specified reliability indicators at maximum efficiency (information transfer rate).

Results. The dependences of reliability indicators (bit and block error coefficients) on the properties of the wireless information transmission channel and the parameters of the control system elements are obtained. Recommendations are given on the use of the results obtained when choosing the settings for the control element (smartphone) and the controlled element (communication processor of the hearing aid). To conduct experimental research, control system models in the MathWorks MatLab Simulink environment were created and tuned.

Conclusions. The studies carried out in the work allow us to calculate and reasonably choose the parameters of the devices of the wireless control system for the given reliability indicators taking into account the error behavior model in the transmission channel and the settings of the associated equipment. This makes it possible to design and implement reliable control systems with specified reliability indicators and maximum information transfer rate.

KEYWORDS: information management systems, reliability, error rate, simulation, Bluetooth.

ABBREVIATIONS

ARQ – Automatic Repeat Request;
AWGN – Additive White Gaussian Noise;
BER – Bit Error Rate;
BLE – Bluetooth Low Energy;
CRC – Cyclic Redundancy Check;
DSP – Digital Signal Processing;
FEC – Forward Error Correction;
FER – Frame Error Rate;
HEC – Hybrid Error Correction;
HV – High quality Voice;
IEEE – Institute of Electrical and Electronics Engineers;
SEQN – Sequence Numbering;
VoIP – Voice over Internet;
Wi-Fi – Wireless Fidelity;
3-D – 3-Dimensional;
av – average value.

NOMENCLATURE

B – number of error bits;
 B_0 – total number of transmitted bits;
 C – speed of light;
 Ch – Wi-Fi channel number;
 d – distance between devices;
 E_s/N_0 – signal to noise ratio;
 f – functional;
 F – frequency;
 F – number of error frames;
 F_0 – total number of transmitted frames;
 k – signal transmission channel number;
 L_0 – attenuation in free space;
 N – availability and settings of other devices operating in the same frequency range;
 P – transmitter power;
 t – simulation time;
 T – type of technology;
 V_p – phase speed;
 α – the actual part of the distribution coefficient;
 β – the imaginary part of the propagation coefficient (phase coefficient);
 γ – distribution coefficient;
 λ – wavelength;
 ρ – interference characteristics;
 ω – angular velocity.

INTRODUCTION

Over the past seventy years, active aging of the society of all developed countries has been observed throughout the world. According to the United Nations, in 1950 the number of people over 60 years old was approximately 202 million. To date, the number of older people is close to one billion people. There is also information about the continuation of the aging trend of society. An increase in the number of elderly people entails an increase in age-related diseases [1]. Also, with the increase in the number of vehicles, with the development of industry, the number of emergencies and,

as a result, the number of injuries sustained by a person are growing [2].

From the above facts, it can be concluded that the replenishment of the functional capabilities and aesthetic appearance of a person is an urgent problem and requires special attention, both from doctors and engineers.

So, for example, with complete or partial lack of hearing, we suggest using an ear implant, which will be an apparatus to restore the ability to perceive sound information [3]. An important role for ensuring the quality functioning of the hearing aid is played by the organization of a high-speed and reliable control channel using specialized software running on a modern smartphone [4, 5]. One of the most significant parameters is the reliability of communication, which is characterized by error rates (bit and block) and depends on the data transfer rate, interference parameters, type of protocol selected, etc. [6].

The object of study was a wireless control channel between a smartphone and a hearing aid, implemented using Bluetooth technology.

The subject of the study is an experimental study of the program model of the wireless control channel in order to determine the characteristics of reliability and quality of information transfer.

The purpose of the work is the formation of practical recommendations on the use of Bluetooth technology in wireless channels for controlling medical devices using an example of a hearing aid.

1 PROBLEM STATEMENT

It was found, that during the operation of the hearing aid, there is a need for its increased functionality [7], such as:

- volume control from a smartphone;
- the ability to answer incoming calls;
- wireless charge;
- several settings modes.

Different settings modes are needed for a more comfortable use of the hearing aid. For example, one of them will be for everyday use, in which the amplifications at different frequency intervals will be individually configured for each patient, based on data obtained from an examination by an audiologist. And the second mode is for use in noisy places. It will be tuned to sufficiently strong noise reduction and speech emphasis. There may also be more modes, depending on the wishes of the patient. Switching between modes, as well as adjusting the volume, will be carried out through a special application in the smartphone.

To solve the problem of choosing wireless technology for the implementation of the control channel with the specified confidence indicators, we perform the mathematical formulation of the problem.

Let there be a combination of communication channel parameters and control system devices: interference characteristics (E_s/N_0), transmitter power (P), distance between devices (d), the presence and settings of other devices operating in the same frequency range (N), type

of technology (T). An urgent task is to determine and study the functional dependencies of reliability (quality) indicators of communication, which are characterized by error coefficients, from these parameters. Each indicator can be represented by a set of functional dependencies of the following form:

$$\text{BER, FER} = f(\rho); d = \text{const}; P = \text{const}. \quad (1)$$

$$\text{BER, FER} = f(\rho); d = \text{const}; N = \text{const}. \quad (2)$$

$$\text{BER, FER} = f(d); E_s/N_0 = \text{const}; P = \text{const}. \quad (3)$$

$$\text{BER, FER} = f(d); E_s/N_0 = \text{const}; T = \text{const}. \quad (4)$$

...

To solve the problems of synthesis of a control system with predetermined reliability indicators, it is necessary to study the dependences of the entered transmission indicators at various values of the communication channel parameters that may occur during the practical implementation of the control system.

Reliability is characterized by the coefficients of bit and block errors. Efficiency is specified by the channel bandwidth efficiency, while the actual task is to maximize it (minimize symbolic redundancy) while observing a given reliability. Therefore, questions are further resolved to determine and study the dependences of reliability and efficiency indicators on the parameters of the transmission medium and control system devices. This makes it possible to select and implement in the elements of control systems ways to ensure a given reliability at the maximum information transfer rate.

Based on the foregoing, an urgent task is to choose the technology for communication between a smartphone and a hearing aid. This article will consider the possibility of using modern wireless technologies. In addition, control system models will be built and analyzed, as well as recommendations for the practical implementation of the control channel and the settings of the control system devices will be given.

2 REVIEW OF THE LITERATURE

In recent years, rehabilitation medicine has made great strides, which mainly occur due to the appearance of bionic prostheses. Bionic prosthetics is the replacement of lost body parts with electronic implants. The creation of such implants allowed the study of biocurrents generated by the human brain [8].

The era of bionics began not so long ago, but it is already difficult to surprise people with an artificial organ or limb, which perform all the necessary functions and almost do not differ in appearance from biological ones. Patients, whose situation was not so long ago regarded as almost hopeless, with the use of bionic devices, it became possible to return to full value and improve the quality of their life.

Bionic prosthetics is an innovative, rapidly developing field of science. Many world scientific studies are aimed at further development of various types of bionic implants, aimed both at patients with various acquired diseases or injuries, and for patients with defects from

birth. Certain successes in this area have been achieved in the development of bionic arms and legs, artificial heart and retina [9].

To restore hearing, there are many different solutions: hearing aids of air or bone conduction, cochlear implants, but they all have their drawbacks, the main of which is a violation of the aesthetic appearance of a person [10].

The solution offered by our regional research group is unique and consists in creating a bionic ear. The main idea is to create a hearing aid using modern components and place it in the auricle implant printed using 3-D technologies from materials that are most similar in appearance and tactile feel to the biological ear.

To implement the settings and use of the hearing aid, it is necessary to calculate the parameters that critically affect the wireless data transfer [11]. Moreover, it is effective to use modern approaches to ensure the reliability of information transmission [12, 13] against the background of the effects of interference of various nature and properties [14, 15].

For the study, a number of modern wireless communication technologies were analyzed, and Bluetooth-LE was selected based on a comparative analysis [16]. Unlike classic Bluetooth, BLE is designed to provide significantly lower power consumption. This allows smartphone applications to interact with BLE devices with more stringent power requirements, such as proximity sensors, heart rate monitors, fitness devices, in our case, a hearing aid. The basis of interaction between devices operating on the basis of the BLE protocol is the client-server architecture. The phone supports a central role, the hearing aid supports the role of a peripheral device, one of two is required to establish a BLE connection [17]. Devices that support only the role of peripheral devices cannot communicate with each other, and two devices cannot communicate with only a central connection. As soon as the telephone and the hearing aid have established a connection, they begin to transmit metadata to each other [18]. Depending on the type of data they transmit, one or the other may act as a server. For example, if the hearing aid wants to transfer data to the telephone, it might make sense for the hearing aid to act as a server. If the hearing aid wants to receive updates from the phone, then it may make sense for the phone to act as a server, which is important in our project, the point-to-point connection type. In this regard, the urgent task is to develop and study a data transfer model between a smartphone and hearing aids via the Bluetooth LE protocol.

3 MATERIALS AND METHODS

As part of the project, it is necessary to study the physical parameters of the signal during transmission from the source to the recipient. Both the processor of the hearing aid and the smartphone can be a source and recipient, since data transmission must be carried out in both directions. The task was to determine the optimal number of frequency channels, the number of which will achieve the target value – the distance, in the presence of

interference. Also, determine the noise immunity of data transmission. The result of the analysis was obtained on the basis of modeling.

The designed communication system was modeled on standard blocks Matlab Simulink, DSP Blockset and the Communications Blockset library [19]. One way to build such a system in Simulink is to start by designing the communication channel and developing it by adding modulations, FEC, etc. testing at every stage. You can develop pairs of components separately, for example, a speech encoder and decoder that can be created and tested in their own model, and then built into the system. The generalized structural diagram of the model is shown in Fig. 1.

When creating the network model, the Bluetooth Voice Transmitter model was taken as the basis (Fig. 2), and its structure was adapted to bring the model closer to the research tasks.

Fig. 2 shows the top level of the complete Bluetooth Voice Simulink model. It includes:

1. Master Transmitter – model of the signal transmitter, i.e. a transmitting device, for example, a smartphone;
2. AWGN – model of a radio channel in which «white noise» operates;

3. Free Space Path Loss – a simulator of a long wireless line, which determines the loss in free space;

4. 802.11b Interferer – model of the interference source in the form of a transmitter operating over the IEEE 802.11b Wi-Fi protocol,

5. Slave Receiver – the receiver of the slave node, ie, the receiving device, for example, another smartphone or Hands-Free unit in the car.

The hierarchical features of Simulink modeling allow you to create large, complex structures that are controlled from subsystems. The discovery of these subsystems shows further levels of detail [19].

Consider the principle of building the main blocks of the model.

The transmitter whose internal circuit is shown in Fig. 3, contains: input of random binary values, coding, buffering, framing, HEC/FEC modulations and frequency hopping.

All real-life continuous communications transmitted in communication systems reflect processes whose main part of the spectrum is concentrated in a finite frequency range. This is explained by the frequency properties of message sources and subscribers (message recipients), which are real physical systems [20].

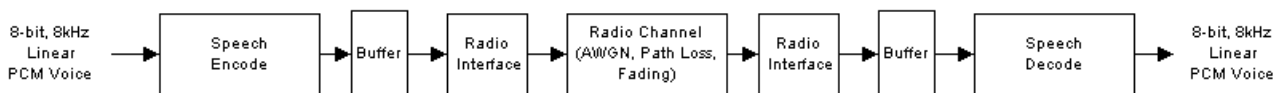


Figure 1 – Generalized block diagram of a Bluetooth control channel model

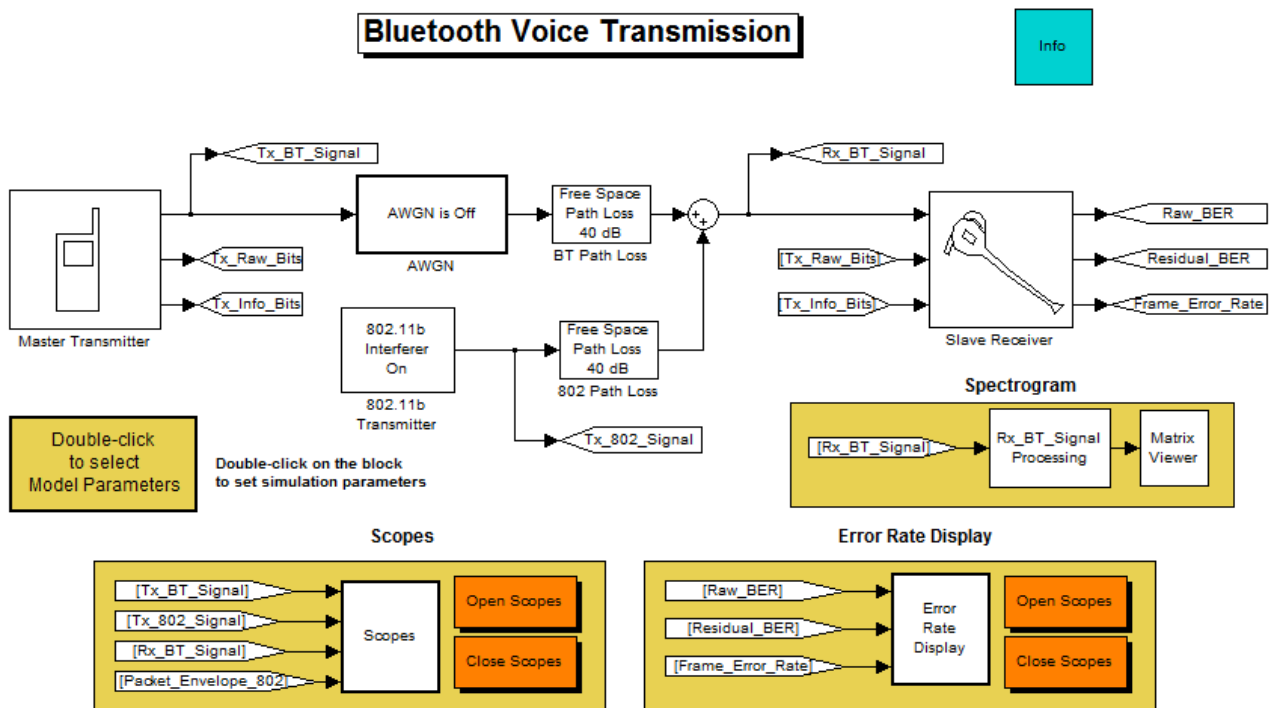


Figure 2 – Bluetooth Voice Transmitter in Simulink

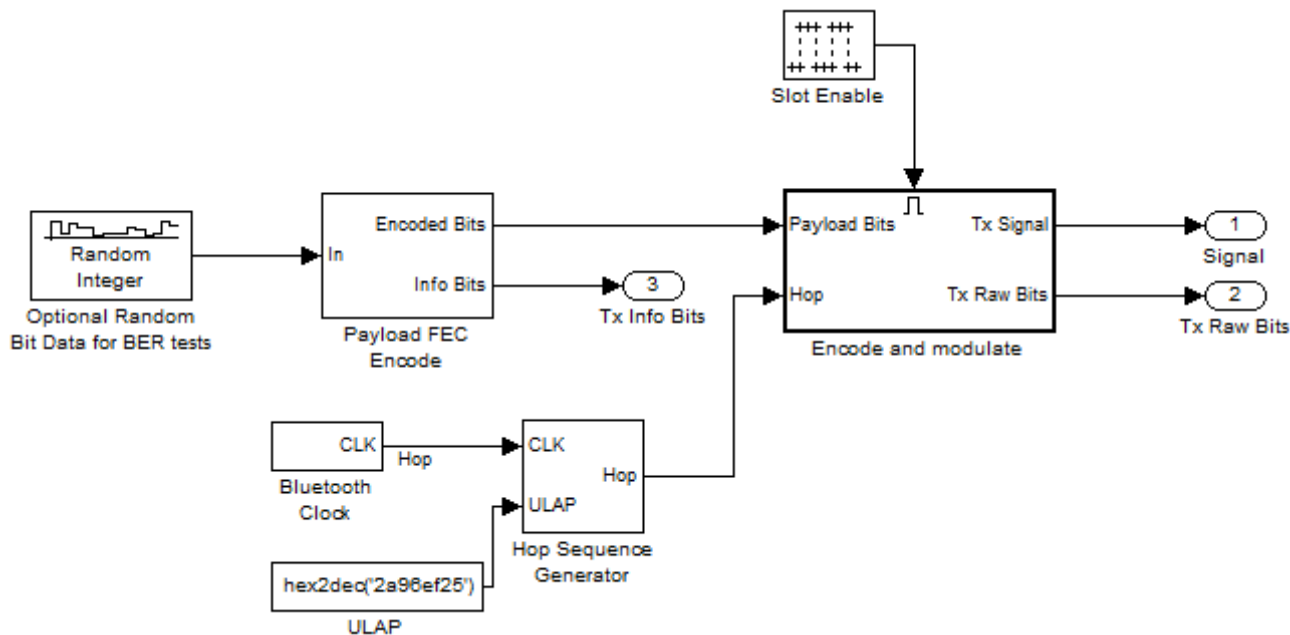


Figure 3 – Internal structure of the signal transmitter unit

Bluetooth technology uses time division duplex (multiplexing). The master device transmits packets at odd time intervals (slots), and the slave device transmits packets at even times. Packages depending on the length can take up to five intervals. Moreover, the channel frequency does not change until the end of the packet transmission [20].

A packet is a format of bits organized into an array that are transmitted over a physical channel. A package consists of an access code, a package header, and user information. The structure of the message packet is shown in Fig. 4.

The access code identifies packets belonging to the same piconet, and is also used for synchronization and request procedures. It includes a preamble (4 bits), a sync word (64 bits), and a frame checksum – 4 bits of a checksum. The header contains information for managing communications and consists of six fields:

1. Address (3 bits) – address of the active element;
2. Type (4 bits) – data type code;
3. FLOW-F (1 bit) – data flow control, shows the readiness of the device to receive;
4. ARQ-A (1 bit) – confirmation of the correct reception;
5. SEQN-S (1 bit) – serves to determine the sequence of packets;
6. Checksum (8 bits) – control checksum.

The final part of the overall package format is user information. It consists of three segments: the header of the user information, the user information itself and the CRC.

The header (8 bits) defines a logical channel, flow control in logical channels, and also has a pointer to the length of user information. User Information CRC (16 bits) – 16 bits of cyclic redundancy code is calculated from the transmitted information, after which it is attached to the information [21].

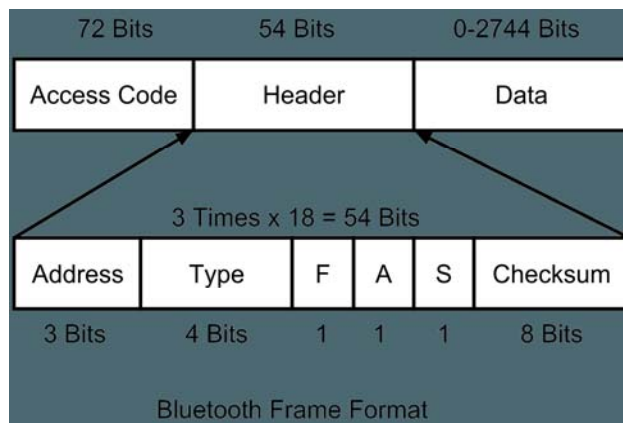


Figure 4 – Bluetooth message packet structure

The AWGN is a simulator of white noise. The Bluetooth Voice Transmitter model allows you to set the signal to noise ratio in the AWGN unit. This ratio is an argument to the function resulting from the study. The menu for setting the parameters of the AWGN block is shown in Fig. 5.

Attenuation in free space is calculated by the formula (5):

$$L_0 = 20 \lg \frac{4\pi d}{\lambda} \text{ [dB]}, \quad (5)$$

L_0 – attenuation; d – attenuation distance; λ – wavelength [19].

The wavelength, as is known, depends on the phase coefficient β (6) (this is the imaginary part of the propagation coefficient $\gamma = \alpha + j\beta$) [21]:

$$\lambda = \frac{2\pi}{\beta}. \quad (6)$$

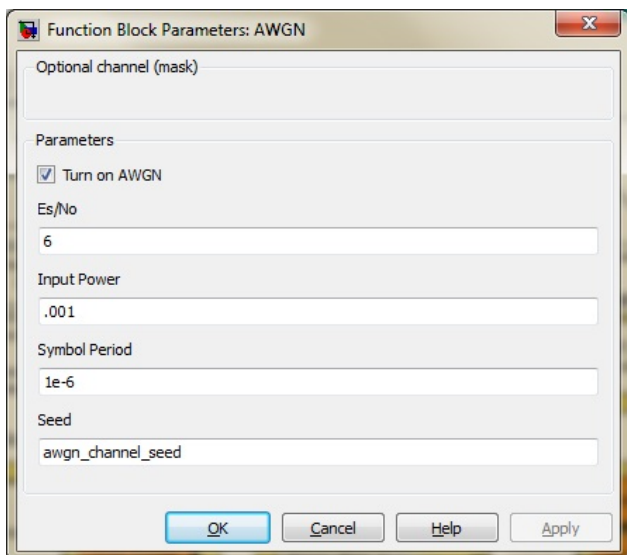


Figure 5 – The menu for setting the parameters of the AWGN block

And the phase velocity also depends on β (7) [21]:

$$V_p = \frac{\omega}{\beta}. \quad (7)$$

In air $V_p = C$ (speed of light $\approx 3 \cdot 10^8$ m/s). So, from (6) and (7) get (8):

$$\beta = \frac{\omega}{V_p} = \frac{\omega}{C} = \frac{2\pi f}{C} \Rightarrow \lambda = \frac{2\pi C}{2\pi f} = \frac{C}{f}. \quad (8)$$

Therefore, the formula for calculating the loss in free space will take the form (9):

$$L_0 = 20 \lg \frac{4\pi d f}{C}. \quad (9)$$

In the considered Bluetooth model, there are two options for setting free space loss during wave propagation:

1. The reference attenuation from 10 to 40 dB, shown in Fig. 6;

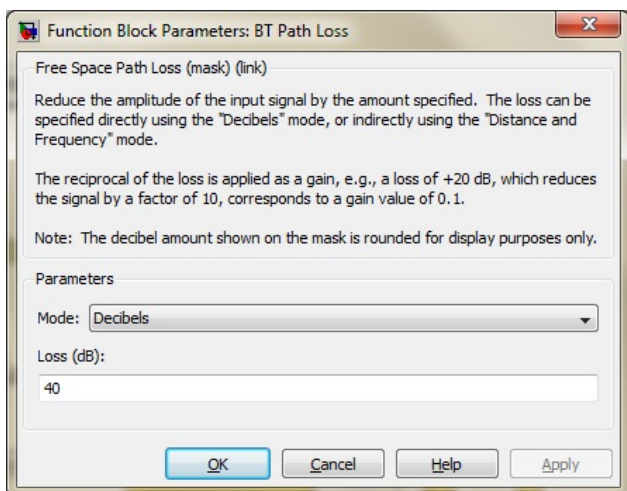


Figure 6 – Window for selecting parameters for setting free space loss in decibels

2. Setting the frequency and transmission distance (with these parameters, the model itself calculates the free space loss in dB) shown in Fig. 7 [21].

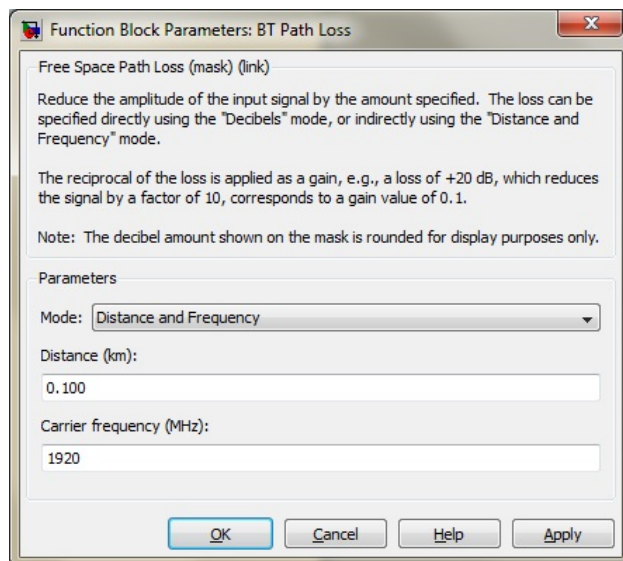


Figure 7 – Parameters selection window for setting the operating frequency and distance

To study the dependence of the error coefficient on the attenuation in the communication line, various parameters were set to set the loss in free space.

To calculate the transmission distance for a given attenuation we solve the logarithmic equation:

$$40 = 20 \lg \frac{4\pi d f}{C}. \quad (10)$$

The calculation formula for determining the transmission range in general is:

$$d = \frac{C \cdot 10^{\frac{L_0}{20}}}{4\pi f}. \quad (11)$$

To calculate the distance d (transmission distance for a given attenuation), we solve the logarithmic equation (11).

For the radio exchange of a Bluetooth device, the frequency range 2400–2483.5 MHz is used. The capacity of this frequency band is 79 subchannels with a bandwidth of 1 MHz. The carrier frequency of the subchannels $f_k = 2402 + k$ (MHz), where $k = 0, \dots, 78$. To reduce the complexity of transceivers, radio channels with binary frequency modulation are used. The coding is simple – the logical unit corresponds to a positive frequency deviation, to zero – negative [22].

On 39th channel, carrier frequency will be equal:

$$f_k = 2402 + 39 = 2441 \text{ MHz}, \quad (12)$$

on 79th channel

$$f_k = 2402 + 78 = 2480 \text{ MHz}. \quad (13)$$

So for channel 39:

1. Attenuation in free space from formula (7) at $d = 2$ m (the maximum distance at which Bluetooth is planned to be used when both network nodes are located by one user), $f_k = 2.441$ GHz, is equal to:

$$L_0 = 20 \lg \frac{4 \cdot \pi \cdot 2.441 \cdot 10^9}{3 \cdot 10^8} = 40.19312 \text{ dB.} \quad (14)$$

2. Transmission distances determined by the formula (11) when $L_0 = 40$ dB, $f_k = 2.441$ GHz:

$$d = \frac{3 \cdot 10^8 \cdot 10^{\frac{40}{20}}}{4 \cdot \pi \cdot 2.441 \cdot 10^9} = 0.978011 \text{ m.} \quad (15)$$

For channel 79:

1. Attenuation in free space from formula (5) at $d = 2$ m (the traditional distance at which Bluetooth works stably in practice) and $f_k = 2.441$ GHz is equal to:

$$L_0 = 20 \lg \frac{4 \cdot \pi \cdot 2.480 \cdot 10^9}{3 \cdot 10^8} = 40.3264 \text{ dB.} \quad (16)$$

2. Transmission distance from formula (11) when $L_0 = 40$ dB и $f_k = 2.441$ GHz:

$$d = \frac{3 \cdot 10^8 \cdot 10^{\frac{40}{20}}}{4 \cdot \pi \cdot 2.480 \cdot 10^9} = 0.963119 \text{ m.} \quad (17)$$

Thus, for 39 and 79 channels, a difference in attenuation of 0.133 dB was obtained, and a transmission distance of 0.015 m.

The simulation result will be the values of the error rate: bit error rate and frame error rate, which are calculated by formulas (18) and (19), respectively [23]:

$$BER = \frac{B}{B_0}, \quad (18)$$

$$FER = \frac{F}{F_0}. \quad (19)$$

4 EXPERIMENTS

Next, the conducted experiments will be described and conclusions made on them.

Parameters of the first experiment: the Wi-Fi unit is turned off, the simulation time $t = 0.5$ (simulated simulation clocks), the interference power is «white noise» Input Power = 0.001 W.

The dependence of the following form is investigated:

$$BER = f(Es/No) | d = \text{const.}$$

Table 1 – BER, FER versus Es/No dependency, for Input Power = 0.001 W

Es/No	$d = 1 \text{ m}$		$d = 5 \text{ m}$		$d = 10 \text{ m}$	
	BER	FER	BER	FER	BER	FER
2	0.2390	1	0.2416	1	0.2438	1
5	0.1345	0.9174	0.1382	0.9009	0.1366	0.8960
8	0.0554	0.0218	0.0529	0.2030	0.0534	0.2256
11	0.0115	0.0075	0.0118	0	0.0121	0
14	0.0006	0	0.0009	0	0.0008	0
17	0	0	0	0	0	0
20	0	0	0	0	0	0

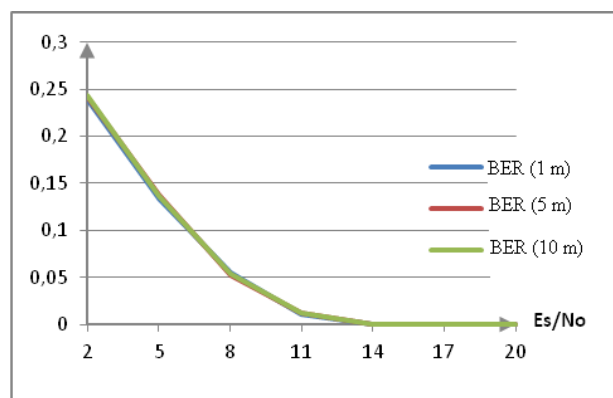
Let's plot the studied dependences (Fig. 8).

The conclusion of Fig. 8: the error rates decrease with increasing signal-to-noise ratio and are not significantly dependent on the distance between the devices. This allows you to provide high-quality communication at the most frequently used in practice, the removal of devices from each other.

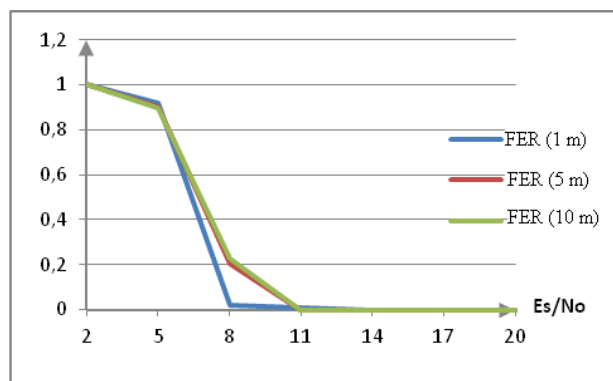
Increase the input noise white power to 0.01 watts and run the simulation again.

Let's plot the studied dependences (Fig. 9).

The conclusion of Fig. 9: the error coefficients, as in the previous simulation, decrease with increasing signal-to-noise ratio and depend insignificantly on the distance between the devices. An increase in the power of the white noise signal entailed an increase in the number of bit and frame errors compared to the previous simulation.



a



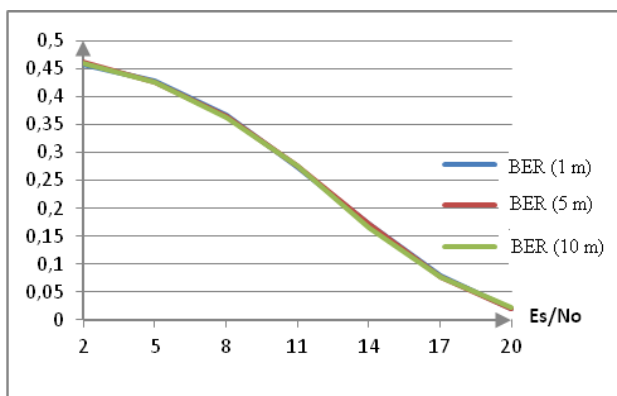
b

Figure 8 – Graph of BER and FER versus Es/No for $d = \text{const}$ and Input Power = 0.001 W

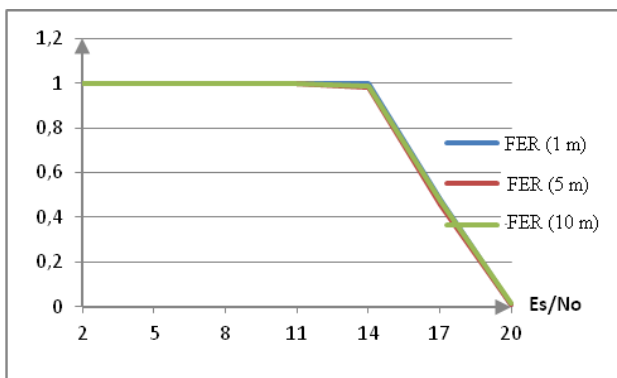
Table 2 – BER and FER versus E_s/No dependency for Input Power = 0.01 W

Es/No	$d = 1\text{ m}$		$d = 5\text{ m}$		$d = 10\text{ m}$	
	BER	FER	BER	FER	BER	FER
2	0.4575	1	0.4610	1	0.4609	1
5	0.4276	1	0.4252	1	0.4261	1
8	0.3668	1	0.3643	1	0.3615	1
11	0.2747	1	0.2763	1	0.2761	1
14	0.1714	1	0.1717	0.9804	0.1658	0.9901
17	0.0797	0.4812	0.0771	0.4586	0.0777	0.4737
20	0.0212	0.0075	0.0206	0.0075	0.0216	0.0150

Once again, increase the white noise power Input Power by 10 times to 0.01 W and perform the simulation again.



a



b

Figure 9 – Graph of BER and FER versus E_s/No at $d = \text{const}$ and Input Power = 0.01 W

Table 3 – BER и FER versus E_s/No dependency for Input Power = 0.1 W

Es/No	$d = 1\text{ m}$		$d = 5\text{ m}$		$d = 10\text{ m}$	
	BER	FER	BER	FER	BER	FER
2	0.4942	1	0.4937	1	0.4931	1
5	0.4918	1	0.4870	1	0.4887	1
8	0.4875	1	0.4805	1	0.4865	1
11	0.4682	1	0.4662	1	0.4687	1
14	0.4382	1	0.4452	1	0.4342	1
17	0.3911	1	0.3891	1	0.3868	1
20	0.3113	1	0.3084	1	0.3079	1

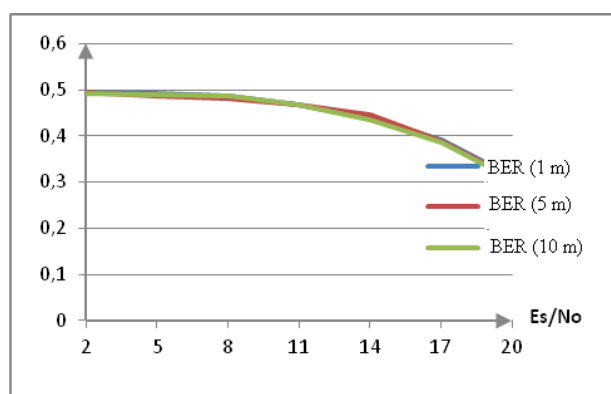
Let's plot the studied dependences (Fig. 10).

The conclusion of fig. 10: as in the previous two cases, the error coefficients decrease with increasing signal-to-noise ratio and depend insignificantly on the distance between the devices. An increase in the power of the white noise signal by another 10 times again entailed an increase in the number of bit and frame errors, and led to the fact that all the received frames at all the studied distances turned out to be erroneous.

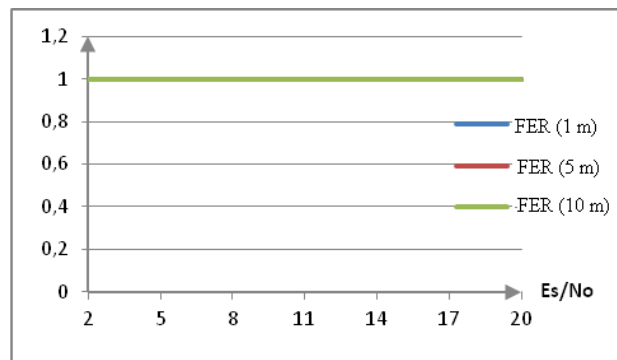
Parameters of the second experiment: signal-to-noise ratio = 10, white noise signal power = 0.001, simulation time $t = 0.5$ (simulated simulation clocks), Wi-Fi signal power = 0.1 W.

The dependence of the following form is investigated:

$$BER = f(\text{Wi-Fi } Ch) | d = \text{const.}$$



a



b

Figure 10 – Graph of BER and FER versus E_s/No at $d = \text{const}$ and Input Power = 0.1 W

Table 4 – BER и FER versus Wi-Fi Ch dependency for Wi-Fi Power = 0.1 W

Es/No	$d = 1\text{ m}$		$d = 5\text{ m}$		$d = 10\text{ m}$	
	BER	FER	BER	FER	BER	FER
No Wi-Fi	0.0206	0.0226	0.0196	0.0075	0.0215	0.0075
0	0.0207	0.0225	0.0402	0.0677	0.0335	0.0226
39	0.0330	0.0301	0.0374	0.0602	0.0362	0.0602
78	0.0277	0.0301	0.0344	0.0376	0.0408	0.0827

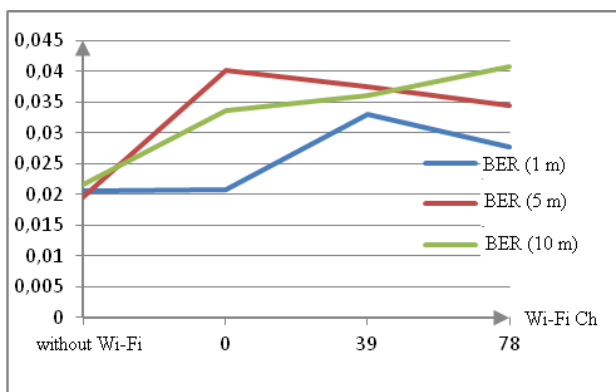
Let's plot the studied dependences (Fig. 11).

The conclusion of Fig. 11: When interference with a Wi-Fi signal of a relatively high power, the error coefficients have a random distribution pattern, which can be explained by the different influence of the signal in the working frequency range of Bluetooth.

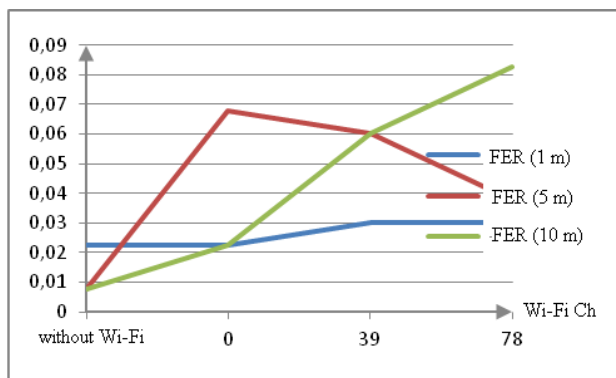
Reduce the power of the Wi-Fi signal to 0.01 and explore.

Table 5 – BER, FER versus Wi-Fi Ch dependency for Wi-Fi Power = 0.01 W

Es/ No	$d = 1\text{ m}$		$d = 5\text{ m}$		$d = 10\text{ m}$	
	BER	FER	BER	FER	BER	FER
No Wi-Fi	0.0206	0.0226	0.0196	0.0075	0.0215	0.0075
0	0.0242	0.0075	0.0366	0.0376	0.0278	0.0451
39	0.0270	0.0451	0.0302	0.0451	0.0429	0.0677
78	0.0248	0.0150	0.0293	0.0526	0.0222	0.0150



a



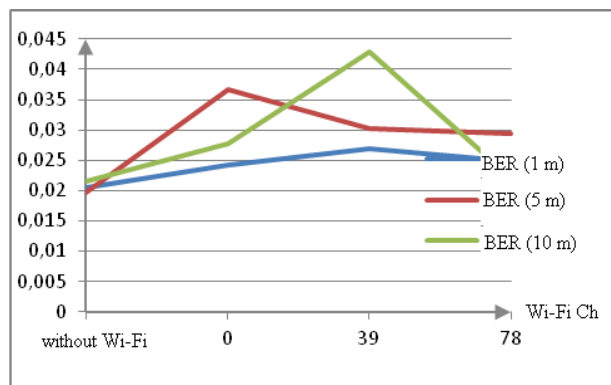
b

Figure 11 – Graph of BER and FER versus Wi-Fi channel for $d = \text{const}$ and Wi-Fi Power = 0.1 W

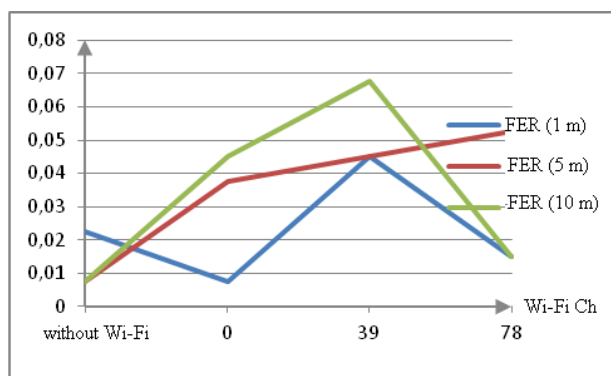
Let's plot the studied dependences (Fig. 12).

The conclusion of Fig. 12: The error rates are mostly extreme when connecting a signal source on a Wi-Fi channel, the closest frequency range to the working range of Bluetooth frequencies.

Once again, we reduce the Wi-Fi signal strength so that it is equal to 0.001, and investigate.



a



b

Figure 12 – Graph of BER and FER versus Wi-Fi channel for $d = \text{const}$ and Wi-Fi Power = 0.01 W

Table 6 – BER, FER versus Wi-Fi Ch dependency for Wi-Fi Power = 0.001 W

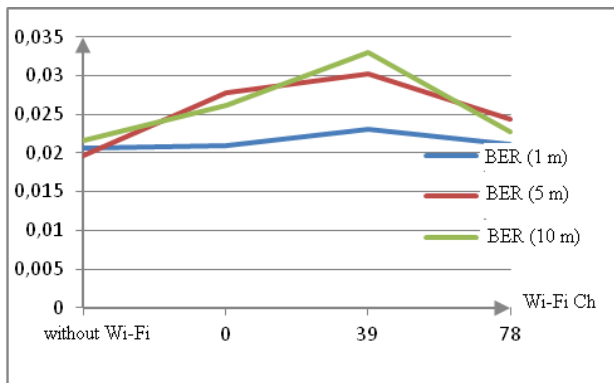
Es/ No	$d = 1\text{ m}$		$d = 5\text{ m}$		$d = 10\text{ m}$	
	BER	FER	BER	FER	BER	FER
No Wi-Fi	0.0206	0.0226	0.0196	0.0075	0.0215	0.0075
0	0.0210	0.0226	0.0278	0.0376	0.0262	0.0226
39	0.0230	0.0376	0.0302	0.0451	0.0330	0.0301
78	0.0211	0.0226	0.0244	0.0376	0.0228	0.0075

Let's plot the studied dependences (Fig. 13).

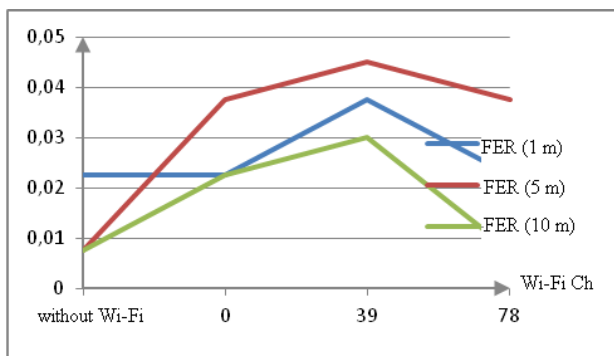
The conclusion of Fig. 13: The error rates are maximum when connecting the signal source on the Wi-Fi channel, the closest frequency range to the working frequency range of Bluetooth.

Third experiment parameters: signal-to-noise ratio = 10, white noise signal power = 0.001, Wi-Fi signal power = 0.01 W, Wi-Fi channel of the Wi-Fi transmitter $Ch = 39$, distance $d = 5\text{ m}$. The dependence of the following form is investigated:

$$\text{BER} = f(t) \mid d = \text{const.}$$



a



b

Figure 13 – Graph of BER and FER versus Wi-Fi channel for $d = \text{const}$ and Wi-Fi Power = 0.001 W

Table 7 – BER versus simulation time t

Simulation time	BER 1	BER 2	BER 3	BER 4	BER _{av}
0.01	0.0191	0.0246	0.0155	0.0228	0.0205
0.05	0.0240	0.0151	0.0246	0.0240	0.02191
0.1	0.0574	0.0447	0.0406	0.0247	0.04185
0.5	0.0356	0.0397	0.0335	0.0315	0.0351
1	0.0309	0.0400	0.0407	0.0273	0.0347
5	0.0358	0.0351	0.0374	0.0366	0.0362
10	0.0387	0.0383	0.0339	0.0374	0.0371

Table 8 – FER versus simulation time t

Simulation time	FER 1	FER 2	FER 3	FER 4	FER _{av}
0.01	0	0	0	0	0
0.05	0	0	0.0769	0	0.0192
0.1	0.1111	0.0741	0.0370	0.0370	0.0648
0.5	0.0677	0.0451	0.0301	0.0451	0.0470
1	0.0375	0.0599	0.0674	0.0300	0.0487
5	0.0533	0.0443	0.0488	0.0503	0.0491
10	0.0516	0.0642	0.0465	0.0533	0.0539

Let's plot the studied dependences (Fig. 14).

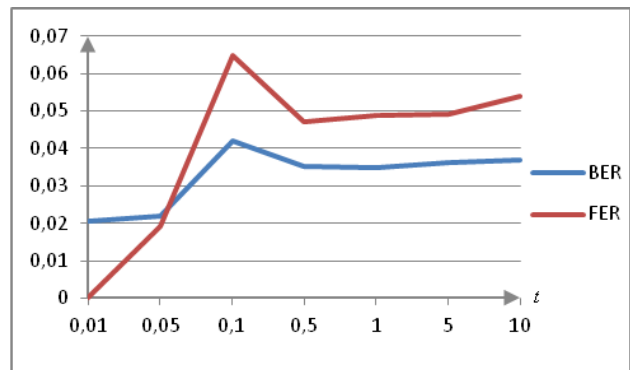


Figure 14 – BER and FER versus simulation time t dependency for $d = \text{const}$

The conclusion of Fig. 14: the graphs show an extremum at 0.1 conditional clock cycles of modeling, but this is a consequence of a random set of transmitted data and the random nature of the interference. When conducting repeated simulations, the extremum may correspond to a different value of the simulation time, or it may be absent altogether (modeling was carried out, but the graphs are not given). There is also averaging and stabilization of the values of bit and block errors with increasing simulation time.

Parameters of the fourth experiment: signal-to-noise ratio = 10, white noise signal power = 0.001, Wi-Fi signal power = 0.01 W, Wi-Fi channel of the Wi-Fi transmitter $Ch = 39$, distance $d = 5 \text{ m}$, simulation time $t = 0.5$ (conditional cycles of modeling). The dependence of the following form is investigated:

$$\text{BER} = f(\text{HV}) \mid d = \text{const.}$$

Table 9 – BER and FER versus Bluetooth transmission packet type

Packet type	BER	FER
HV1	0.0296	0.0324
HV2	0.0395	0.0498
HV3	0.0368	0.0602

Let's plot the studied dependences (Fig. 15).

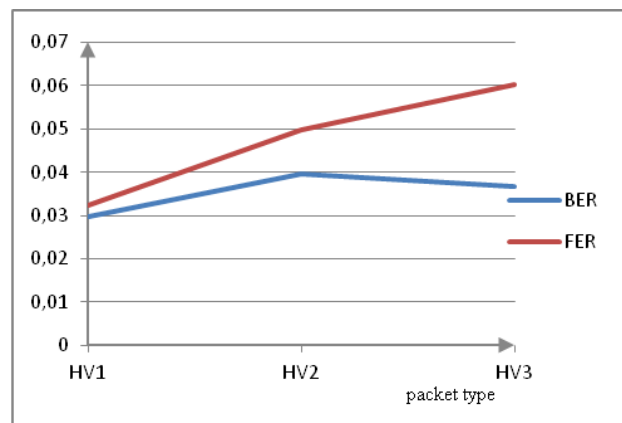


Figure 15 – BER and FER versus packet type dependency for $d = \text{const}$

The conclusion of Fig. 15: the smallest value of bit and frame errors is observed when using a packet of type HV1. Also noticeable is the change in the difference between the values of bit and block errors when transmitting different packets. This is due to the peculiarities of data structures in packets of different types.

5 RESULTS

As a result of the theoretical studies, dependencies are obtained that are essential for the design of wireless control channels, reliability indicators for interference characteristics, signal power, distance, the presence of sources of interfering signals of other technologies, etc. Analytical dependencies were determined, the influence of parameters on the integral indicators of reliability was investigated, their graphic illustrations were obtained, and an analysis of their nature was made.

The simulation circuit model of a wireless control system that implements the communication technology under study in the modeling package MatLab Simulink is investigated. With its help, experimental studies were carried out. They confirmed the introduced mathematical position. Another purpose of the studied models is the tasks of analysis and synthesis of control systems with given indicators of reliability and efficiency.

6 DISCUSSION

The several wireless technologies are selected for comparison: Bluetooth, Wi-Fi, ZigBee [7]. The protocol specifications similar in technical characteristics (IEEE 802.11b and IEEE 802.15.4) are chosen. A comparison is based on the results of data presented in scientific publications and open information sources [14, 16, 17, 20]. The comparison was carried out according to several basic characteristics: the signal-to-noise ratio, power consumption, transmission speed, noise immunity. The conclusion of comparison is Bluetooth technology has large advantages. Therefore, it was chosen for modeling and research in order to obtain effective characteristics for configuration in the developed device.

The studies of reliability indicators of the wireless control channel between the smartphone and the hearing aid and the resulting dependencies were used for choosing the configurations and settings of the control system elements. For given technical conditions – the signal-to-noise ratio, protocol type, packet size, distance between devices, the presence or absence of other signal sources, one can select such settings to ensure maximum reliability of information speed for high efficiency of using the resources of the developed control system.

CONCLUSIONS

The results of the study confirmed that to solve the problem – “Wireless interaction between the smartphone and hearing system”, Bluetooth LE technology meets all the requirements for frequency and energy characteristics.

The scientific novelty of the presented results lies in the fact that, as a result of analytical and experimental

research, the nature of the influence of interference parameters, transmission channel characteristics and network configuration on the appearance, structure and level of errors in the operation of the wireless control system was revealed.

The practical significance of the results of the work lies in the implementation of simulation circuit models of a wireless control system using the selected wireless transmission technology. They allow you to perform studies of reliability indicators and evaluate the effectiveness of using reliability indicators for various interference characteristics, device parameters and system configurations. Models are an effective toolkit for the analysis and synthesis of control systems with predetermined indicators of reliability and energy efficiency.

Prospects for further research are expected in the development of algorithmic and software tools for calculating and comparative analysis of the characteristics of wireless communication technologies to select an effective data transfer platform for control systems.

ACKNOWLEDGEMENTS

A regional research group uniting employees of the Perm National Research Polytechnic University and Perm State Medical University carried out the presented studies as part of joint research and development work. E. A. Wagner and Moscow State Medical and Dental University. A. I. Evdokimova. The results are intended for hardware and software implementation of the developed hearing aid. This work was supported by a grant from the Perm Region. Documents filed for an invention patent.

REFERENCES

1. Zeng F. Challenges in Improving Cochlear Implant Performance and Accessibility, *IEEE Transactions on Biomedical Engineering*, 2017, Vol. 64, No. 8, pp. 1662–1664. DOI: 10.1109/TBME.2017.2718939.
2. Arefjeva N. A., Saveljeva E. E. Objective diagnosis of frequency thresholds of hearing in young children, *Russian otorhinolaryngology*, 2016, No. 6, pp. 17–26.
3. Parris S., Torlak M., Kehtarnavaz N. Real-time implementation of cochlear implant speech processing pipeline on smartphones, *IEEE Engineering in Medicine and Biology Society, 36th Annual International Conference, Chicago, 26–30 August 2014, proceedings*. Chicago, IEEE, 2014, pp. 886–889. DOI: 10.1109/EMBC.2014.6943733.
4. Hansen J. H. L. CCI-MOBILE: Design and Evaluation of a Cochlear Implant and Hearing Aid Research Platform for Speech Scientists and Engineers, *IEEE-EMBS Biomedical & Health Informatics (BHI), International Conference, Chicago, 19–22 May 2019, proceedings*. Chicago, IEEE, 2019, pp. 1–4. DOI: 10.1109/BHI.2019.8834652.
5. Audio Processor Specification [Electronic resource]. Access mode: www.onsemi.com/pub/Collateral/E7150-D.PDF.
6. IC MEMS AUDIO SENSOR 3RHLGA [Electronic resource]. Access mode: <https://www.digikey.com/product-detail/en/stmicroelectronics/MP23AB02BTR/497-14947-1-ND/4988080>.

7. Bajdarov A. A., Gurko V. A. et al. The use of digital systems to eliminate the physiological problems of perception of sound information, *Automated control systems and information technology, Russian scientific and technical conference, Perm, 22–24 May 2019, proceedings*. Perm, PNRPU, 2019, Vol. 2, pp. 115–119.
8. Buyens W., Van Dijk B., Wouters J. et al. A harmonic/percussive sound separation based music pre-processing scheme for cochlear implant users, *21st European Signal Processing Conference (EUSIPCO 2013, International Conference, Marrakech, 9–13 September 2013, proceedings*. Marrakech, IEEE, 2013, pp. 1–5.
9. Chen K., Li Q., Li W. et al. Three-dimensional finite element modeling of Cochlear implant induced electrical current flows, *Computational Intelligence for Measurement Systems and Applications, IEEE International Conference, Hong Kong, 11–13 May 2009, proceedings*. Hong Kong, IEEE, 2009. – P. 5–7. DOI: 10.1109/CIMSA.2009.5069907.
10. Derbel A., Ghorbel M., Samet M. et al. New algorithm of bionic wavelet transform to DSP-implementation for cochlear implant, *IEEE Mediterranean Electrotechnical Conference, 16th International Conference, Yasmine Hammamet, 25–28 March 2012 : proceedings*. Yasmine Hammamet, IEEE, 2012, pp. 1025–1029. DOI: 10.1109/MELCON.2012.6196602.
11. Freyman V. Methods and algorithms of soft decoding for signals within information transmission channels between control systems elements, *Radio electronics, Computer science, Control*, 2018, No. 4 (47), pp. 226–235. DOI: 10.15588/1607-3274-2018-4-22.
12. Viterbi A. J., Omura J. K. Principles of digital communication and coding. New York, McGraw-Hill, 2009. 584 p.
13. Freyman V. Research of the Reed-Solomon codes characteristic for realization within control systems devices, *Radio electronics, Computer science, Control*, 2019, № 3 (50), pp. 143–151. DOI: 10.15588/1607-3274-2019-3-16.
14. Sklar B. Digital communications. Fundamentals and applications : second edition. New Jersey, Prentice Hall, 2001, 1079 p.
15. Freyman V. Analysis of the transmission reliability between control systems devices when errors are packaged, *Radio electronics, Computer science, Control*, 2019, № 4 (51), pp. 223–231. DOI: 10.15588/1607-3274-2019-4-21.
16. Bluetooth technology [Electronic resource]. Access mode: <https://www.bluetooth.com>.
17. Multipath based channel simulation with using IEEE 802.11b [Electronic resource]. Access mode: <https://m.habr.com/ru/post/122544/>.
18. On the allocation of radio frequency bands to short-range devices [Electronic resource]. Access mode: <http://docs.cntd.ru/document/902048009>.
19. MatLab [Electronic resource]. Access mode: <http://matlab.ru>.
20. Shmalko A. V. Digital communication networks: the basics of planning and building. Moscow, Eco-Trends, 2001, 284 p.
21. The principle of rapid spasmodic interference [Electronic resource]. Access mode: <http://www.gaw.ru/html.cgi/txt/doc/Wireless/bluetooth/2.htm>.
22. Kuznetsova T. A., Repp P. V. Interference immunity of wireless telecommunication systems in the conditions of channel interference [Electronic resource]. Access mode: <http://www.fundamental-research.ru/ru/article/view?id=33740>.
23. Freyman V., Bezukladnikov I. Research and application of noise stability providing methods at information and control systems, *2017 IEEE Conference of Russian young researchers in electrical and electronic engineering.: International Conference, Saint-Petersburg, 1–3 February 2017, proceedings*. Saint-Petersburg, Saint-Petersburg Electrotechnical University «LETI», 2017, pp. 831–837.

Received 27.05.2020.
Accepted 20.08.2020.

УДК 621.391:004.052

РОЗРОБКА І ДОСЛІДЖЕННЯ БЕЗДРОТОВИХ СИСТЕМ УПРАВЛІННЯ ПРИЛАДОМ «БІОНІЧНЕ ВУХО»

Арутюнов С. Д. – доктор наук, завідувач кафедру «Професорська стоматологічних захворювань», Московський державний медико-стоматологічний університет імені А. І. Євдокимова, м. Москва, Росія.

Асташина Н. Б. – доктор наук, завідувач кафедри «Ортопедична стоматологія», Пермський державний медичний університет ім. академіка О. О. Вагнера, м. Перм, Росія.

Байдаров А. А. – кандидат технічних наук, проректор з інформаційних технологій та інноваційного розвитку, Пермський державний медичний університет ім. академіка О. О. Вагнера, м. Перм, Росія.

Гурко В. А. – старший викладач кафедри «Автоматика і телемеханіка», Пермський державний технічний університет, м. Перм, Росія.

Майоров П. В. – студент кафедри «Автоматика і телемеханіка», Пермський державний технічний університет, м. Перм, Росія.

Фрейдман В. І. – доктор технічних наук, професор кафедри «Автоматика і телемеханіка», Пермський державний технічний університет, м. Перм, Росія.

Южаков А. А. – доктор технічних наук, завідувач кафедри «Автоматика і телемеханіка», Пермський державний технічний університет, м. Перм, Росія.

АНОТАЦІЯ

Актуальність. Проведено аналіз показників достовірності передачі інформації між елементами бездротової системи управління – керуючим пристроєм, реалізованим на смартфоні, і комунікаційним процесором створюваного слухового апарату. Об'єктом дослідження є показники достовірності бездротового каналу управління. Предмет дослідження – теоретичний і експериментальний аналіз залежностей показників достовірності проєктованої бездротової системи управління.

Мета роботи. Визначення і дослідження аналітичних і експериментальних залежностей показників достовірності передачі по бездротовому каналу управління від властивостей каналу зв'язку і налаштувань елементів системи, формування рекомендацій налаштування параметрів елементів системи управління.

Методи. Використано елементи теорії надійності для визначення залежностей коефіцієнтів бітових і блокових помилок від властивостей каналу зв'язку конфігурації елементів системи управління. Отримано аналітичні співвідношення для визначення показників достовірності передачі з урахуванням можливих спотворень сигналів. Досліджено залежності показників достовірності від параметрів елементів системи управління, наведені ілюструють приклади. Розроблено імітаційні схемотехнічні моделі системи управління з обраної бездротової технології Bluetooth. Проведено експериментальні дослідження, на підставі отриманих даних зроблено висновки і запропоновані рекомендації щодо вибору конфігурації системи управління з метою забезпечення заданих показників достовірності при максимальній ефективності (інформаційної швидкості передачі).

Результати. Отримані залежності показників достовірності (коефіцієнти бітових і блокових помилок) від властивостей бездротового каналу передачі інформації і параметрів елементів системи управління. Надано рекомендації щодо використання отриманих результатів при виборі налаштувань керуючого елемента (смартфона) і керованого елемента (комунікаційного процесора слухового апарату). Для проведення експериментальних досліджень створено і налаштовано моделі системи управління в середовищі MathWorks MatLab Simulink.

Висновки. Проведені в роботі дослідження дозволяють розрахувати і обґрунтовано вибрати параметри пристроїв бездротової системи управління для заданих показників достовірності з урахуванням моделі поведінки помилок в каналі передачі і налаштувань супутнього обладнання. Це дає можливість проектувати і реалізувати надійні системи управління із заданими показниками достовірності та максимальної інформаційної швидкості передачі.

КЛЮЧОВІ СЛОВА: інформаційно-керуючі системи, достовірність, коефіцієнт помилок, модель, Bluetooth.

УДК 621.391:004.052

РАЗРАБОТКА И ИССЛЕДОВАНИЕ БЕСПРОВОДНОЙ СИСТЕМЫ УПРАВЛЕНИЯ УСТРОЙСТВОМ «БИОНИЧЕСКОЕ УХО»

Арутюнов С. Д. – доктор наук, заведующий кафедрой «Профилактика стоматологических заболеваний», Московский государственный медико-стоматологический университет имени А. И. Евдокимова, г. Москва, Россия.

Асташина Н. Б. – доктор наук, заведующая кафедрой «Ортопедическая стоматология», Пермский государственный медицинский университет им. академика Е. А. Вагнера, г. Пермь, Россия.

Байдаров А. А. – кандидат технических наук, проректор по информационным технологиям и инновационному развитию, Пермский государственный медицинский университет им. академика Е. А. Вагнера, г. Пермь, Россия.

Гурко В. А. – старший преподаватель кафедры «Автоматика и телемеханика», Пермский национальный исследовательский политехнический университет, г. Пермь, Россия.

Майоров П. В. – студент кафедры «Автоматика и телемеханика», Пермский национальный исследовательский политехнический университет, г. Пермь, Россия.

Фрейман В. И. – доктор технических наук, профессор кафедры «Автоматика и телемеханика», Пермский национальный исследовательский политехнический университет, г. Пермь, Россия.

Южаков А. А. – доктор технических наук, заведующий кафедрой «Автоматика и телемеханика», Пермский национальный исследовательский политехнический университет, г. Пермь, Россия.

АННОТАЦИЯ

Актуальность. Проведен анализ показателей достоверности передачи информации между элементами беспроводной систем управления – управляющим устройством, реализованным на смартфоне, и коммуникационным процессором разрабатываемого слухового аппарата. Объектом исследования являются показатели достоверности беспроводного канала управления. Предмет исследования – теоретический и экспериментальный анализ зависимостей показателей достоверности проектируемой беспроводной системы управления.

Цель работы. Определение и исследование аналитических и экспериментальных зависимостей показателей достоверности передачи по беспроводному каналу управления от свойств канала связи и настроек элементов системы, формирование рекомендаций настройке параметров элементов системы управления.

Методы. Используются элементы теории надежности для определения зависимостей коэффициентов битовых и блоковых ошибок от свойств канала связи конфигурации элементов системы управления. Получены аналитические соотношения для определения показателей достоверности передачи с учетом возможных искажений сигналов. Исследованы зависимости показателей достоверности от параметров элементов системы управления, приведены иллюстрирующие примеры. Разработаны имитационные схемотехнические модели системы управления с выбранной беспроводной технологией Bluetooth. Проведены экспериментальные исследования, на основании полученных данных сделаны выводы и предложены рекомендации по выбору конфигураций системы управления с целью обеспечения заданных показателей достоверности при максимальной эффективности (информационной скорости передачи).

Результаты. Получены зависимости показателей достоверности (коэффициенты битовых и блоковых ошибок) от свойств беспроводного канала передачи информации и параметров элементов системы управления. Даны рекомендации по использованию полученных результатов при выборе настроек управляющего элемента (смартфона) и управляемого элемента (коммуникационного процессора слухового аппарата). Для проведения экспериментальных исследований созданы и настроены модели системы управления в среде MathWorks MatLab Simulink.

Выводы. Проведенные в работе исследования позволяют рассчитать и обоснованно выбрать параметры устройств беспроводной системы управления для заданных показателей достоверности с учетом модели поведения ошибок в канале

передачи и настроек сопутствующего оборудования. Это дает возможность проектировать и реализовать надежные системы управления с заданными показателями достоверности и максимальной информационной скоростью передачи.

КЛЮЧЕВЫЕ СЛОВА: информационно-управляющие системы, достоверность, коэффициент ошибок, модель, Bluetooth.

ЛИТЕРАТУРА / LITERATURA

1. Zeng F. Challenges in Improving Cochlear Implant Performance and Accessibility / F. Zeng // IEEE Transactions on Biomedical Engineering. – 2017. – Vol. 64, № 8. – P. 1662–1664. DOI: 10.1109/TBME.2017.2718939.
2. Арефьева Н. А. Объективная диагностика частотных порогов слуха у детей раннего возраста / Н. А. Арефьева, Е. Е. Савельева // Российская оториноларингология. – 2016. – № 6. – С. 17–26.
3. Parris S. Real-time implementation of cochlear implant speech processing pipeline on smartphones / S. Parris, M. Torlak, N. Kehtarnavaz // IEEE Engineering in Medicine and Biology Society : 36th Annual International Conference, Chicago, 26–30 August 2014 : proceedings. – Chicago : IEEE, 2014. – P. 886–889. DOI: 10.1109/EMBC.2014.6943733.
4. Hansen J. H. L. CCI-MOBILE: Design and Evaluation of a Cochlear Implant and Hearing Aid Research Platform for Speech Scientists and Engineers / J. H. L. Hansen // IEEE-EMBS Biomedical & Health Informatics (BHI) : International Conference, Chicago, 19–22 May 2019 : proceedings. – Chicago : IEEE, 2019. – P. 1–4. DOI: 10.1109/BHI.2019.8834652.
5. Спецификация аудио процессора [Электронный ресурс]. – Режим доступа: www.onsemi.com/pub/Collateral/E7150-D.PDF.
6. IC MEMS AUDIO SENSOR 3RHLGA. [Электронный ресурс]. – Режим доступа: <https://www.digikey.com/product-detail/en/stmicroelectronics/MP23AB02BTR/497-14947-1-ND/4988080>.
7. Применение цифровых систем для устранения физиологических проблем восприятия звуковой информации / [А. А. Байдаров, В. А. Гурко, П. В. Майоров, О. А. Марков] // Автоматизированные системы управления и информационные технологии : Всероссийская научно-техническая конференция, Пермь, 22–24 мая 2019 : материалы конференции. – Пермь : ПНИПУ, 2019. – С. 115–119.
8. A harmonic/percussive sound separation based music pre-processing scheme for cochlear implant users / [W. Buyens, B. van Dijk, J. Wouters et al.] // 21st European Signal Processing Conference (EUSIPCO 2013) : International Conference, Marrakech, 9–13 September 2013 : proceedings. – Marrakech : IEEE, 2013. – P. 1–5.
9. Three-dimensional finite element modeling of Cochlear implant induced electrical current flows / K. Chen, Q. Li, W. Li et al.] // Computational Intelligence for Measurement Systems and Applications : IEEE International Conference, Hong Kong, 11–13 May 2009 : proceedings. – Hong Kong : IEEE, 2009. – P. 5–7. DOI: 10.1109/CIMSA.2009.5069907.
10. New algorithm of bionic wavelet transform to DSP-implementation for cochlear implant / [A. Derbel, M. Ghorbel, M. Samet et al.] // IEEE Mediterranean Electrotechnical Conference : 16th International Conference, Yasmine Hammamet, 25–28 March 2012 : proceedings. – Yasmine Hammamet : IEEE, 2012. – P. 1025–1029. DOI: 10.1109/MELCON.2012.6196602.
11. Freyman V. Methods and algorithms of soft decoding for signals within information transmission channels between control systems elements / V. Freyman // Radio electronics, Computer science, Control. – 2018. – № 4 (47). – P. 226–235. DOI: 10.15588/1607-3274-2018-4-22.
12. Viterbi A. J. Principles of digital communication and coding / A. J. Viterbi, J. K. Omura. – New York : McGraw-Hill, 2009. – 584 p.
13. Freyman V. Research of the Reed-Solomon codes characteristic for realization within control systems devices / V. Freyman // Radio electronics, Computer science, Control. – 2019. – № 3 (50). – P. 143–151. DOI: 10.15588/1607-3274-2019-3-16.
14. Sklar B. Digital communications. Fundamentals and applications : second edition / B. Sklar. – New Jersey : Prentice Hall, 2001. – 1079 p.
15. Freyman V. Analysis of the transmission reliability between control systems devices when errors are packaged / V. Freyman // Radio electronics, Computer science, Control. – 2019. – № 4 (51). – P. 223–231. DOI: 10.15588/1607-3274-2019-4-21.
16. Технология Bluetooth [Электронный ресурс]. – Режим доступа: <https://www.bluetooth.com>.
17. Имитация многолучевого канала на основе IEEE 802.11b [Электронный ресурс]. – Режим доступа: <https://m.habr.com/ru/post/122544/>.
18. О выделении полос радиочастот устройствам малого радиуса действия [Электронный ресурс]. – Режим доступа: <http://docs.cntd.ru/document/902048009>.
19. MatLab [Электронный ресурс]. – Режим доступа: <http://matlab.ru>.
20. Шмалько А. В. Цифровые сети связи: основы планирования и построения / А. В. Шмалько. – М. : Экотрендз, 2001. – 284 с.
21. Принцип быстрого скачкообразного изменения ИНТЕРФЕРЕНЦИИ [Электронный ресурс]. – Режим доступа: <http://www.gaw.ru/html.cgi/txt/doc/Wireless/bluetooth/2.htm>.
22. Кузнецова Т. А. Помехоустойчивость систем беспроводной электросвязи в условиях канальной интерференции [Электронный ресурс] / Т. А. Кузнецова, П. В. Репп. – Режим доступа: <https://www.fundamental-research.ru/ru/article/view?id=33740>.
23. Freyman V. Research and application of noise stability providing methods at information and control systems / V. Freyman, I. Bezukladnikov // 2017 IEEE Conference of Russian young researchers in electrical and electronic engineering : International Conference, Saint-Petersburg, 1–3 February 2017 : proceedings. – Saint-Petersburg : Saint-Petersburg Electrotechnical University «LETI», 2017. – P. 831–837.

Наукове видання

**Радіоелектроніка,
інформатика,
управління**

№ 3/2020

Науковий журнал

Головний редактор – д-р техн. наук С. О. Субботін

Заст. головного редактора – д-р техн. наук Д. М. Піза

Комп'ютерне моделювання та верстання
Редактор англійських текстів

С. В. Зуб
С. О. Субботін

Оригінал-макет підготовлено у редакційно-видавничому відділі НУ «Запорізька політехніка»

Свідоцтво про державну реєстрацію
КВ № 24220-14060 ПР від 19.11.2019.

*Підписано до друку 08.10.2020. Формат 60×84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 24,18.
Тираж 300 прим. Зам. № 1072.*

69063, м. Запоріжжя, НУ «Запорізька політехніка», друкарня, вул. Жуковського, 64

Свідоцтво суб'єкта видавничої справи
ДК № 6952 від 22.10.2019.