UDC 004.056.55, 004.942

PROPERTIES OF GENERATORS OF PSEUDO-RANDOM SEQUENCES CONSTRUCTED USING FUZZY LOGIC AND TWO-DIMENSIONAL CHAOTIC SYSTEMS

Kushnir M. Ya. – PhD, Associate Professor, Associate Professor of the Department of Radio Engineering and Information Security. Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine.

Kosovan Hr. V. – PhD, Assistant of the Department of Radio Engineering and Information Security, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine.

Kroyalo P. M. – Postgraduate student, Department of Radio Engineering and Information Security, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine.

ABSTRACT

Context. The problem of generating pseudo-random sequences of bits using the rules of fuzzy logic and two-dimensional chaotic systems is considered.

Objective. Pseudo-random sequences generators built using two-dimensional chaotic systems and fuzzy logic. The purpose of the work is to develop and implement pseudo-random bit sequences generators based on the rules of fuzzy logic and two-dimensional chaotic systems and to evaluate the statistical characteristics of the generated sequences using statistical tests of National Institute of Standards and Technology.

Method. A method for generating pseudo-random bit sequences is proposed, which allows form bit sequences with characteristics that meet the requirements of secure communication systems and cryptographic protection of information based on the rules of fuzzy logic and two-dimensional chaotic systems. In the process of studying the operation of generators, histograms of the distribution of output values were constructed, which allows to clearly determine whether the entire range of output values of the two-dimensional system could be used to generate pseudo-random bit sequence or only part of it. A study of the statistical characteristics of the generated sequences using a set of statistical tests was also performed.

Results. Bit sequences formed using fuzzy logic rules and two-dimensional chaotic systems can be used to transmit information in secure communication systems.

Results. The proposed generators were implemented in software, histogram analysis and evaluation of compliance with the criteria for a set of statistical tests of National Institute of Standards and Technology.

Conclusions. The experiments confirmed the ability of the proposed generators to generate bit sequences with good statistical characteristics, which allows them to be recommended for use in practice in solving problems of cryptographic protection of information and secure transmission of information over open communication channels. Prospects for further research may be to create cryptographic methods of information protection based on the proposed pseudo-random bit sequences generators, the implementation of secure communication systems.

KEYWORDS: generator, chaos, two-dimensional system, pseudo-random sequence, fuzzy logic, statistical tests.

ABBREVIATIONS

FFT is a fast Fourier transform;

NIST is a statistical tests suite of National Institute of Standards and Technology;

PFMM-CLM is a parallel fuzzy multimodule chaotic logistic mapping;

PRB is a pseudo-random bits;

PRS is a pseudo-random sequence.

NOMENCLATURE

x_0 is an initial condition of the two-dimensional Hénon system;

y_0 is an initial condition of the two-dimensional Hénon system;

x_{n+1} is an output value of the two-dimensional Hénon system;

y_{n+1} is an output value of the two-dimensional Hénon system;

a is a control parameter of the two-dimensional Hénon system;

b is a control parameter of the two-dimensional Hénon system;

c_0 is an initial condition of the two-dimensional Lozi system;

d_0 is an initial condition of the two-dimensional Lozi system;

c_{n+1} is an output value of the two-dimensional Lozi system;

d_{n+1} is an output value of the two-dimensional Lozi system;

α is a control parameter of the two-dimensional Lozi system;

β is a control parameter of the two-dimensional Lozi system;

p_0 is an initial condition of the two-dimensional cross-chaotic system;

r_0 is an initial condition of the two-dimensional cross-chaotic system;

p_{n+1} is an output value of the two-dimensional cross-chaotic system;

DOI 10.15588/1607-3274-2022-1-5

r_{n+1} is an output value of the two-dimensional cross-chaotic system;

μ is a control parameter of the two-dimensional cross-chaotic system;

k is a control parameter of the two-dimensional cross-chaotic system;

n is a number of iterations of chaotic system;

$x_{\min}$ is a minimum of range of output values of the two-dimensional Hénon system;

$x_{\max}$ is a maximum of range of output values of the two-dimensional Hénon system;

P_{value} is a criterion for passing the statistical test NIST.

INTRODUCTION

A method of generating pseudo-random sequence (PRS) bits using multidimensional chaotic systems and fuzzy logic rules for the formation of pseudo-random bit sequences with their further verification for compliance with the criteria of statistical tests suite of National Institute of Standards and Technology (NIST) is suggested in this article [1–8]. A number of multidimensional chaotic systems, such as two-dimensional Hénon, Lozi maps, and cross-chaotic maps, are used as mathematical functions for the formation of initial values.

Fuzzy logic in the sense of deterministic chaos is a section of mathematical logic designed to solve the problem of fuzzy decision making by assigning a certain bit value to a fuzzy range of initial values of a chaotic system to obtain the most accurate result possible [9–14]. Fuzzy logic is designed to solve the problem of generating bits by considering all available information and making the best possible decision from the generated initial value of the chaotic system. To verify the effectiveness of this method of generating PRS, the latter should be tested for compliance with the criteria of NIST statistical tests, which will confirm the effectiveness of encoders and cryptographic methods based on such generators for processing, transmitting or storing confidential information [15–19].

A large number of different PRS bit generators are known from the literature that both use threshold methods to generate bit sequences and generate sequences by converting a decimal value into a bit representation [6–8]. We suggest to use the rules of fuzzy logic to form pseudo-random bit sequences.

The object of study is the process of pseudo-random bit sequence generation using two-dimensional chaotic systems and fuzzy logic.

The subject of study is the combination of chaos theory and fuzzy logic rules to form a new approach to creating secure data transmission systems.

The purpose of the work is to develop and implement PRS bit generators based on the rules of fuzzy logic in two-dimensional chaotic systems and to evaluate the statistical characteristics of the generated sequences using statistical tests NIST.

1 PROBLEM STATEMENT

To generate PRS bits, we selected three two-dimensional chaotic mappings using fuzzy logic, namely the Hénon (1), Lozi (2) maps, and cross-chaotic (3) maps [1, 3].

$$\begin{aligned} x_{n+1} &= y_{n+1} - ax_n^2, \\ y_{n+1} &= bx_n^2, \end{aligned} \tag{1}$$

where $x_0 \in (-1;1)$ and $y_0 \in (-0.4;0.4)$ are the initial states of Hénon chaotic systems, $a \in (0;2]$, $b \in (-0.5;0.5]$ are control parameters.

$$\begin{aligned} c_{n+1} &= 1 - \alpha|c_n| + d_n, \\ d_{n+1} &= \beta c_n, \end{aligned} \tag{2}$$

where $c_0 \in (-2;2)$ and $d_0 \in (-2;2)$ are the initial states of chaotic systems, $a \in (1.3;1.8)$ and $b \in (0.3;0.6)$ are control parameters.

$$\begin{aligned} p_{n+1} &= 1 - \mu r_n^2, \\ r_{n+1} &= \cos\left(k \cos^{-1} p_n\right), \end{aligned} \tag{3}$$

where $p_0 \in (-1;1)$ and $y_0 \in (-1;1)$ – are the initial states of chaotic systems, $\mu \in (1,4;2]$ and $k \in (0,3)$ are control parameters.

Depending on how the control parameter of the chaotic system is selected a different range of initial values is obtained, and the formation of bit sequences will be done in a different way. Therefore, in order to be able to form a bit sequence, it is necessary to adapt the rules of fuzzy logic to make them suitable for a bit sequence formation.

Histograms of distribution of initial values of two-dimensional chaotic systems, as well as results of the NIST statistical tests will serve as the criteria of the formed sequences estimation.

2 REVIEW OF THE LITERATURE

Chaos theory is used in numerous applications, namely in cryptography, secure communications, technology, physics, economics, robotics, control and many others [1, 3, 5, 12]. Chaotic systems are deterministic ones with high sensitivity to initial conditions and changes in control parameters and are therefore constitute an excellent basis for effective modeling of complex natural phenomena. These features allow using the chaotic systems to build secure communication systems.

Due to the above characteristics of chaotic systems, there is a constant demand for the introducing new appli-

DOI 10.15588/1607-3274-2022-1-5

cations of chaotic systems in secure communication systems. Usually new applications are implemented by either modifying the existing chaotic system, or by slightly changing the equations describing chaotic systems, or adding another equation to the system and increasing its dimension, or proposing a new application of an already well-studied chaotic system.

Logistic map [7] is one of the most well-known one-dimensional discrete time chaotic systems and one of the most heavily modified chaotic systems [7, 8, 12, 15]. The map has only one parameter and a simple structure, which makes it suitable for many applications. Many modifications of the classical logistic map have been proposed in the literature. One of such modifications is the use of a fuzzy triangular number to change the behavior of the logistic map. The idea of passing the logistic map values through a fuzzy number is mathematically simple, but it leads to a significant improvement in the behavior of chaotic map.

Fuzzy logic and a fuzzy sets themselves are a large field of research and have found their application in technology. Specifically, in dynamical systems fuzzy sets are combined with chaotic systems and form the so-called fuzzy dynamical systems [9–14]. In one of the proposed modifications of the logistic map, its values at each iteration are passed through a triangular fuzzy number, which is a simple linear function that takes values in the interval [0, 1]. The resulting map demonstrates a more unpredictable behavior associated with chaos compared to the classical map, and reaches a higher value for its Lyapunov exponent. In addition, to demonstrate the applicability of the map in chaos-related software applications, the problems of generating pseudo-random values are described in [17–19], and the process of image encryption based on such systems is presented in [19]. It can be seen that a sequence of bits formed from a modified mapping using a simple rule passes all 15 tests of the NIST statistical test suite [15,18]. Further, the generated bit sequence is used to implement the image encryption process, and the resulting encrypted image is analyzed for security using methods such as histogram analysis, correlation and information entropy.

It should be noted that the approach of combining fuzzy logic and chaotic systems can be easily applied to any chaotic system and further modified by considering different types of fuzzy numbers, such as trapezoidal, Gaussian, quadratic, exponential ones or combinations thereof. Thus, the combination of fuzzy logic and chaos is quite promising for the development of new PRS bit generators and their application in cryptography and secure communication systems.

3 MATERIALS AND METHODS

To implement the PRS bit generator, we used the following fuzzy logic rule:

1. First, we divide the range of initial values of each of the chaotic systems into 10 intervals.

2. Each of these intervals is divisible by 25 subintervals except the last one; it is divisible by 30.

3. The size of each of the intervals is determined depending on the selected values of the control parameters of chaotic systems, and they will differ from each other for both systems.

First, we selected four two-dimensional chaotic maps, such as Hénon (1), Lozi (2), and a cross-chaotic map (3) to test the efficiency of such generators.

Fuzzy logic rule for a two-dimensional Hénon map with the values of the control parameter a = 1.40 and b = 0.3035 and the range of change of output values [–1.297; 1.276] is as follows:

If the input = –1.297 – –1.0397, the output = 0–25
If the input = –1.0397– –0.7824, the output = 26–50
If the input = –0.7824 – –0.5251, the output = 51–75
If the input = –0.5251– –0.2678, the output = 76–100
If the input = –0.2678 – –0.0105, the output =101–125
If the input = –0.0105 – 0.2468, the output = 126–150
If the input = 0.2468 – 0.5041, the output = 151–175
If the input = 0.5041 – 0.7614, the output = 176–200
If the input = 0.7614 – 1.0187, the output = 201–225
If the input = 1.0187 – 1.276, the output = 226–255.

Similarly, the original ranges of Lozi maps and the cross-chaotic map are broken and PRS bits are formed. Fig. 1 shows a block diagram of a PRS bit generator using fuzzy logic and two-dimensional chaotic systems.

4 EXPERIMENTS

In the process of studying the statistical characteristics of bit sequences, PRS bits were formed separately for each two-dimensional map with different initial conditions and control parameters. Their initial conditions and control parameters, under which the best results of statistical tests were obtained, are presented in Table 1. In addition, since the system is very sensitive to the values of initial conditions and control parameters, it is necessary to choose the values of control parameters so that the range of initial values of chaotic systems is fully completed.

5 RESULTS

To check whether the whole range of initial values is really completed, it is necessary to build a histogram of the initial values distribution. Since the fuzzy logic rule used for generating bit sequences is divided into 256 intervals, the volume of the histogram will also be confined to 256 intervals. Fig. 2 presents a histogram of the initial values distribution of the Hénon map. The range of output values was from $x_{\min} = -6.3722$ to $x_{\max} = 6.3699$, which was divided into 256 intervals. Here, the abscissa axis shows the division of the range of the initial values into 256 intervals, and the ordinate axis – the number of values that fall into the corresponding interval.

DOI 10.15588/1607-3274-2022-1-5

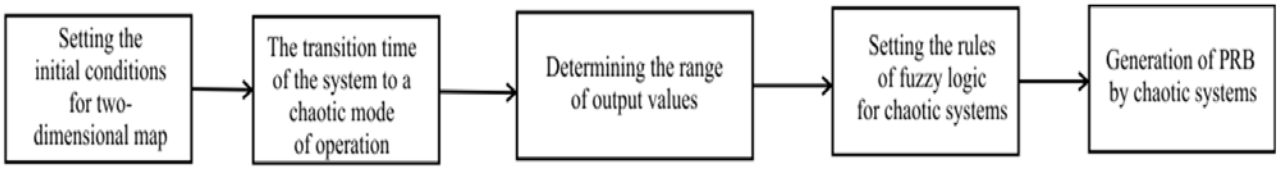


Figure 1 – Block diagram of the PRS bit generator using fuzzy logic and two-dimensional chaotic system

Table 1 – Values of initial conditions and control parameters of two-dimensional chaotic systems

Two-dimensional chaotic map	Henon	Lozi	Cross-chaotic
Initial conditions	x_0 =0.254	c_0 =0.173	p_0 =0.324
	y_0 =0.321	d_0 =0.255	r_0 =0.651
Control parameters	a=0.0413	α =1.6113	μ =2.81
	b=0.99991	β =0.5202	k =7.73

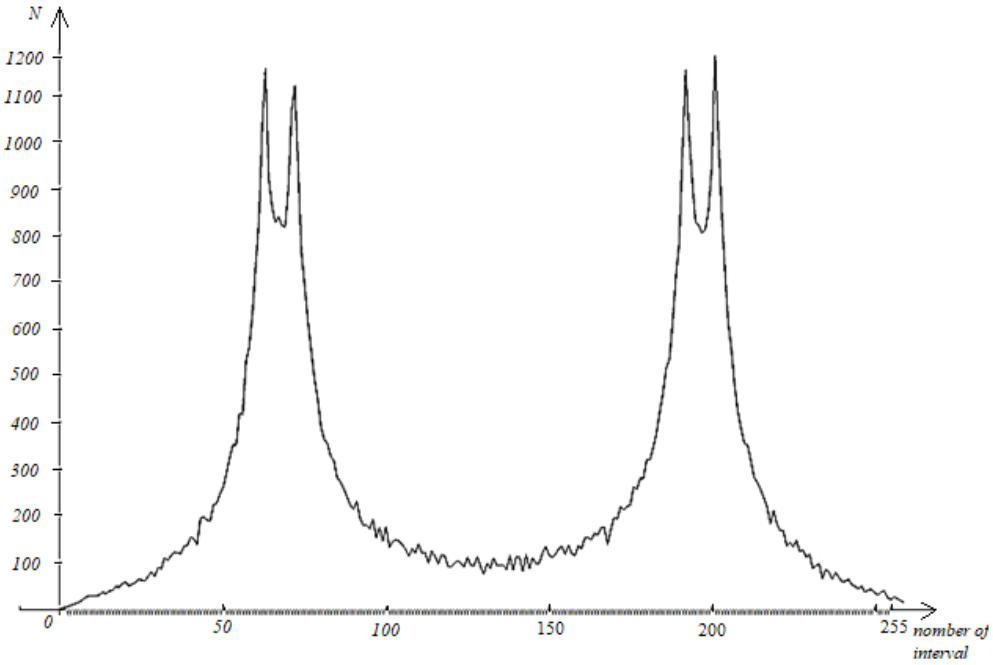


Figure 2 – Histogram of the distribution of the output values of the Henon mapping for 65,000 iterations

It can be seen from the obtained histogram that two areas predominate in the number of values that fall there. The greater number of values falling in a certain area in terms of statistical research or application in cryptography is rather considered a disadvantage. Therefore, to generate PRS bits, it is advisable to use not the entire range of initial values, but only a part of it with uniform distribution of the number of values falling into it.

Fig. 3 presents a histogram of the initial values distribution of the Lozi map. The range of output values was from $x_{\min} = -1.2236$ to $x_{\max} = 1.38$ and was also divided into 256 intervals. It follows from the obtained histogram that the distribution of the initial values is almost uniform, and this enables to generate PRS bits. Uniformity of distribution also determines the use of PRS bit generator based on fuzzy logic in cryptographic and secure communication systems.

The histogram of the initial values distribution for two-dimensional the cross-chaotic map is presented in Fig. 4. The range of output values was from $x_{\min} = -1.1$ to $x_{\max} = 0.96$ and was divided into 256 intervals.

It can be seen from the obtained histogram that, similarly to the case of the Hénon map, there are two areas predominating in the number of values that fall there. Therefore, to generate PRS bits, it is advisable to use not the entire range of the initial values, but only a part of it with the best uniform distribution.

The results of studying the PRS bits for compliance with the criteria of statistical tests formed by the Hénon, Lozi and cross-chaotic maps are presented in Tables 2, 3 and 4.

DOI 10.15588/1607-3274-2022-1-5

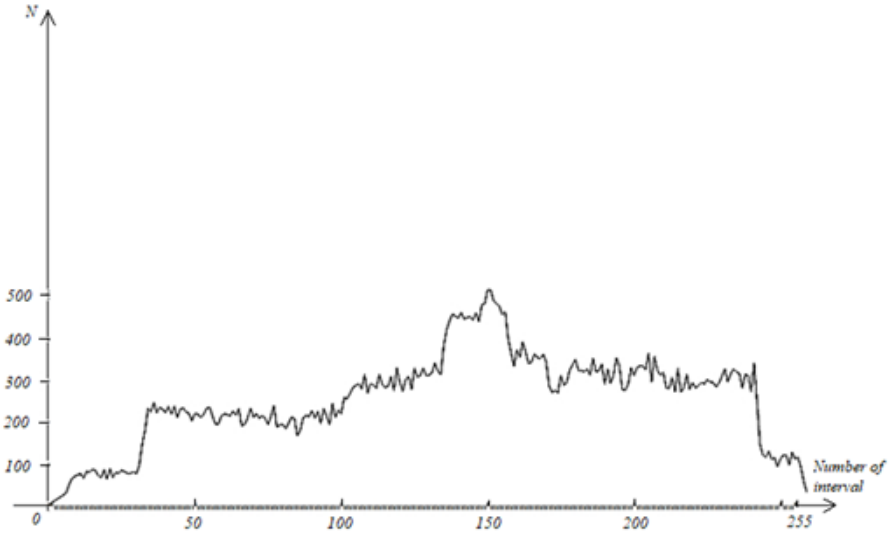


Figure 3 – Histogram of the distribution of the output values of the Lozi mapping for 65,000 iterations

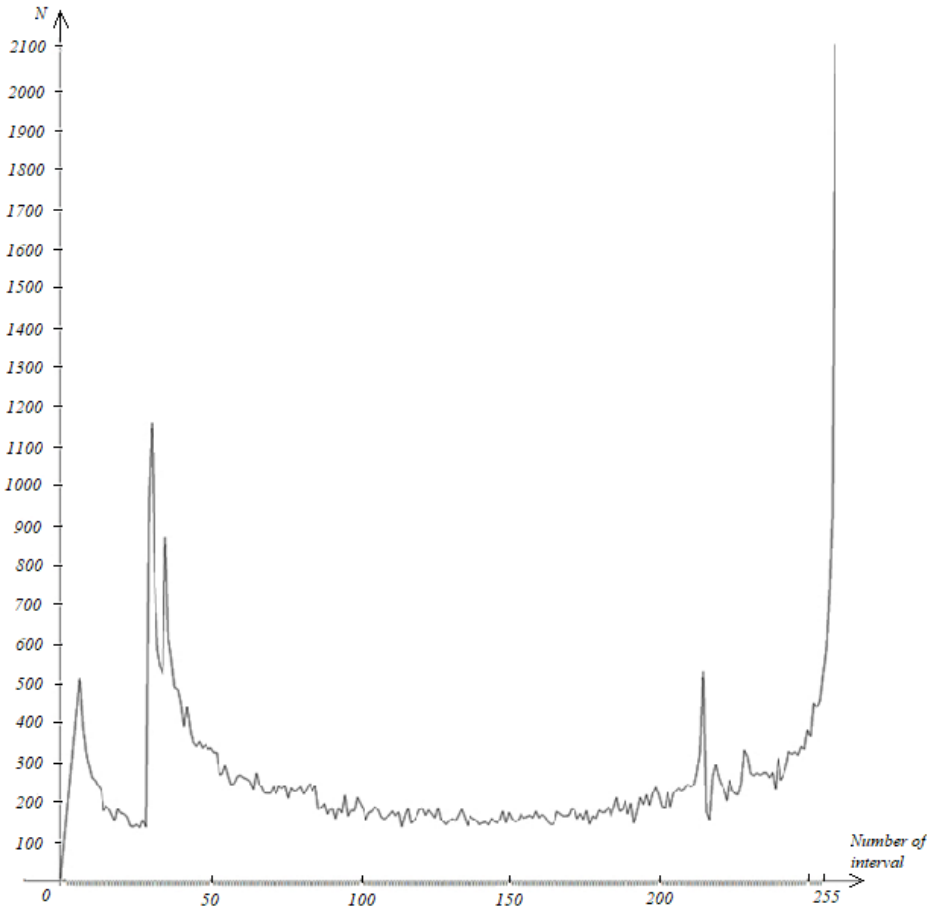


Figure 4 – Histogram of the distribution of the output values of two-dimensional cross-chaotic mapping for 65,000 iterations

DOI 10.15588/1607-3274-2022-1-5

Table 2 – Test results of the generated sequence formed by the Henon mapping

Statistical test type	The obtained P_{value}	Proportion
Frequency	0.991	0.980
Block Frequency	0.992	1,0
Runs	0.872	0.962
Longest Run	0.895	0.941
Rank	0.992	0.854
FFT	0.990	0.268
Non-Overlapping Template	0,383827	0.980
Overlapping Template	0.987	0.252
Universal	0.963	0.106
Linear Complexity	0.971	1.0
Serial	0.987	0.670
Approximate Entropy	0.989	0.959
Cumulative Sums	0.108791	1.0
Random Excursions	0.203	0.643
Random Excursions Variant	0.213	1.0

Table 3 – Test results of the generated sequence formed by the Lozzi mapping

Statistical test type	The obtained P_{value}	Proportion
Frequency	0.258	1.0
Block Frequency	0.687	1.0
Runs	0.697	0.960
Longest Run	0.253	0.960
Rank	0.799	0.66
FFT	0.547	0.91
Non-Overlapping Template	0.350	1.0
Overlapping Template	0.358	0.180
Universal	0.451	0.960
Linear Complexity	0.366	0.970
Serial	0.783	1.0
Approximate Entropy	0.687	0.980
Cumulative Sums	0.316	1.0
Random Excursions	0.857	0.960
Random Excursions Variant	0.751	0.857

Table 4 – Test results of the generated sequence formed by cross-chaotic mapping

Statistical test type	The obtained P_{value}	Proportion
Frequency	0.138	0.980
Block Frequency	0.113	0.970
Runs	0.380	1.0
Longest Run	0.575	0.960
Rank	0.789	0.990
FFT	0.474	1.0
Non-Overlapping Template	0.321	0.980
Overlapping Template	0.574	0.990
Universal	0.525	0.960
Linear Complexity	0.883	1.0
Serial	0.116	0.960
Approximate Entropy	0.233	0.980
Cumulative Sums	0.178	0.83
Random Excursions	0.037	1.0
Random Excursions Variant	0.745	0.980

6 DISCUSSION

We also compared the results with the results of other studies. In [20], a method for generating PRB sequences using fuzzy logic rules and based on chaotic one-dimensional mappings is proposed. The three most well-known one-dimensional mappings were used in the study, namely logistic, square and cubic. As a result of the inspection, it was found that the PRB generated by such mappings meet the conditions of the tests from the NIST set in part. Therefore, it is not desirable to use only one one-dimensional chaotic mapping to form bit sequences using fuzzy logic rules. To solve this problem, in the same work [20], it was proposed to implement a PVP bit generator using two one-dimensional chaotic systems, namely logistic and cubic mappings. As a result, much better results were obtained, namely, the generated sequences correspond to most of the tests from the NIST set.

In [21], one-dimensional logistic mapping was modified using fuzzy triangular numbers. The result is a new modified logistics mapping. Then this mapping was used to generate pseudo-random bits, which gave high positive

DOI 10.15588/1607-3274-2022-1-5

results. Pseudo-random bits were created by comparing the obtained number with the threshold value selected at 0.5. The value of bit 1 was generated if the number is greater than or equal to the threshold, and the value of bit 0 was obtained otherwise.

A set of statistical tests from the National Institute of Standards and Technology NIST 800-22 was also used to verify that the generated sequence was pseudorandom. The obtained results showed that the sequence generated by the modified logistic mapping passes all tests.

In addition, a new parallel fuzzy multimodule chaotic logistic mapping (PFMM-CLM) was proposed in [22]. In the process of research, logistic mapping was used several times with changed control parameters. In this case, fuzzy set theory is used as a fuzzy logic selector to generate pseudo-random bit sequences. As a result of modeling and performance analysis of the proposed pseudo-random bit generator based on PFMM-CLM, high chaotic properties were obtained, such as a reliable bifurcation diagram and a high value of the Lyapunov exponent. Checking the compliance of statistical tests showed that the sequences generated by such a generator are completely satisfactory to all tests.

As a result of analysis and comparison of all considered results it was found that our proposed pseudo-random bit generator has improved statistical properties in comparison with PVP bit generators based on one-dimensional chaotic systems. In addition, it also has a number of advantages, namely:

– the use of two-dimensional display increases the number of initial conditions and control parameters, and, as a consequence, improves the security of information transmission systems;

– does not require additional modifications;

– does not require multiple use of the same display with different values of control parameters and, as a result, our proposed generator will be fast enough.

CONCLUSIONS

The scientific novelty. The method for generating PRS bit sequences using fuzzy logic rules and based on two-dimensional chaotic maps is proposed in this article. Since two-dimensional maps are very sensitive to the values of control parameters, it was first checked whether all the intervals formed by the rules of fuzzy logic are attended by the initial values of chaotic systems. To check that, the histograms of the initial values distribution were built, and the parts with the most uniform distribution were selected from them to form the PRS bits. After obtaining the best histograms, the pseudo-random bit sequences were generated and further verified for compliance with the NIST test criteria.

The practical significance. The sequences verification was performed both for each of the equations of two-dimensional systems separately and after superimposing the initial values using the XOR operation. Due to verification it was found that, when generating sequences by the Hénon map, the sequence formed by the variable y corresponded better to the conditions of the statistical tests. For the sequence formed by the Lozi map, it was the sequence formed by the first equation of the system, and for the cross-chaotic map the first equation shows the best results. PRSs formed in this way satisfy the conditions of the tests from the NIST suite.

Prospects for further research. PRS bits generated using fuzzy logic rules and two-dimensional chaotic systems can be used to develop methods for encrypting information based on them and to create secure telecommunications systems.

DOI 10.15588/1607-3274-2022-1-5

ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Yuriy Fedkovych Chernivtsi national university “Methods of forming signal structures and information processes of software and hardware interaction of broadband telecommunication systems and the Internet of Things” (state registration number 0121U 112870).

REFERENCES

1. Kocarev L. Chaos-based cryptography: A brief overview, *IEEE Circuits and Systems Maganize*, 2001, Vol. 1, pp. 6–21. DOI:10.1109/7384.963463.
2. Semenko A., Kushnir N., Bokla N., Kosovan Hr. Features of creating based on chaos pseudo-random sequences. Modern Problems of Radio Enginiring, Telecommunications, and Computer Science, *XIIth International Conference TCSET' February 20–24 2018: proceedings*. Lviv-Slavsco, Ukraine. 2018, pp. 338–342. DOI:10.20535/2411-2976.22018.
3. Mira C. and all. Chaotic dynamics in two-dimensional noninvertible maps, *World Scientific Series on Nonlinear Science,* 1996, *Series A*, Vol. 20, pp. 185–337. https://doi.org/10.1142/2252.
4. Hénaff S., Taralova I., Lozi R. Dynamical Analysis of a new statistically highly performant deterministic function for chaotic signals generation, *International Conf. on Physics and Control (PhysCon): proceedings*. Catania, Sicily, September 2009, P. 10. HAL Id: hal-00623064.
5. Strogatz S.H. Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering. CRC Press, Boca Raton, FL, USA, 2018, P. 532. ISBN 9780813349107.
6. Huang X., Liu L., Li X., Yu M., Wu Z. New Pseudorandom Bit Generator Based on Mixing Three-Dimensional Chen Chaotic System with a Chaotic Tactics, *Complexity,* 2019, № 44, pp. 1–9. https://doi.org/10.1155/2019/6567198.
7. Wang Y., Liu Z., Ma J., He H. A pseudorandom number generator based on piecewise logistic map, *Nonlinear Dyn.* 2016, No. 83, pp. 2373–2391. https://doi.org/10.1007/s11071-015-2488-0.
8. Murillo-Escobar M., Cruz-Hernández C., Cardoza-Avendaño L., Méndez-Ramírez R. A novel pseudorandom number generator based on pseudorandomly enhanced logistic map, *Nonlinear Dyn.* 2017, No. 87, pp. 407–425. https://doi.org/10.1007/s11071-016-3051-3.
9. Zimmermann H.J. Fuzzy Set Theory – And Its Applications. Springer Science & Business Media. *Berlin. Germany,* 2011, Vol. 21, 525 p. DOI: 10.1007/978-94-015-8702-0.
10. Chakraverty S., Sahoo D. M., Mahato N. R. Concepts of Soft Computing: Fuzzy and ANN with Programming. Springer: Berlin/Heidelberg, Germany, 2019, 198 p. DOI 10.1007/978-981-13-7430-2.
11. Hanss M. Applied Fuzzy Arithmetic: An Introduction with Engineering Applications. Springer, Berlin/Heidelberg, Germany, 2005, 270 p. DOI: 10.1007/b138914.

12. Li Z., Zhang X. On Fuzzy Logic and Chaos Theory: from an Engineering Perspective. In Fuzzy Logic, *A Spectrum of Theoretical & Practical Issues; Springer*. Berlin/Heidelberg, Germany, 2007. pp. 79–97. ISSN: 1434-9922.
13. Porto M., Amato P. A fuzzy approach for modeling chaotic dynamics with assigned properties, *Ninth IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2000: proceedings.* San Antonio, TX, USA, 7–10 May 2000, Vol. 1, pp. 435–440. DOI: 10.1109/FUZZY.2000.838699.
14. Stefanini L., Sorini L., Guerra M. L. Simulation of fuzzy dynamical systems using the LU-representation of fuzzy numbers, *Chaos Solitons Fractals,* 2006, No. 29, pp. 638–652. https://doi.org/10.1016/j.chaos.2005.08.096.
15. Patidar V., Sud K. K., Pareek N. K. A pseudo random bit generator based on chaotic logistic map and its statistical testing, *Informatica,* 2009, No. 33, pp. 441–452.
16. Stojanovski T., Kocarev L. Chaos-based random number generators-part I: analysis [cryptography], *IEEE Trans. Circuits Syst. I Fundam. Theory Appl*., 2001, No. 48, pp. 281–288. DOI: 10.1109/81.915385.
17. François M., Grosges T., Barchiesi D., Erra R. Pseudo-random number generator based on mixing of three chaotic maps, *Commun. Nonlinear Sci. Numer. Simul.,* No. 19, pp. 887–895. https://doi.org/10.1007/978-3-319-06089-7_16.
18. Rukhin A., Soto J., Nechvatal J., Smid M., Barker E. A statistical Test Suite for Random and Pseudorandom Number Generators for cryptographic Applications, *Technical Report; Booz-Allen and Hamilton Inc.* Mclean, VA, USA, 2001. DOI: 10.3390/sym12081202.
19. Alvarez G., Li S. Some basic cryptographic requirements for chaos-based cryptosystems, *Int. J. Bifurc. Chaos,* 2006, No. 16, pp. 2129–2151. DOI: 10.1142/S0218127406015970.
20. Kushnir M., Kosovan Hr., Kroialo P., Komarnytskyy A. Encryption of the Images on the Basis of Two Chaotic Systems with the use of Fuzzy Logic, *15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET 2020: proceedings.* Lviv-Slavsco, Ukraine. February 25–29, 2020, pp. 610–613. DOI: 10.1109/TCSET49122.2020.235504.
21. Moysis L., Volos Ch., Jafari S. et al. Modification of the Logistic Map Using Fuzzy Numbers with Application to Pseudorandom Number Generation and Image Encryption, *Entropy*, 2020, Vol. 22, 474 p. DOI: 10.3390/e22040474.
22. Gad M., Hagras E., Soliman H. et al. A New Parallel Fuzzy Multi Modular Chaotic Logistic Map for Image Encryption, *The International Arab Journal of Information Technology*. March 2021, Vol. 18, No. 2, pp. 227–236. https://doi.org/10.34028/iajit/18/2/12.

Received 21.11.2021.
Accepted 27.01.2022.

УДК 004.056.55, 004.942

ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ ГЕНЕРАТОРІВ ПСЕВДОВИПАДКОВИХ ПОСЛІДОВНОСТЕЙ, ПОБУДОВАНИХ ІЗ ВИКОРИСТАННЯМ НЕЧІТКОЇ ЛОГІКИ ТА ДВОВИМІРНИХ ХАОТИЧНИХ СИСТЕМ

Кушнір М. Я. – канд. фіз.-мат. наук, доцент, доцент кафедри радіотехніки та інформаційної безпеки Чернівецького національного університету імені Юрія Федьковича, Чернівці, Україна.

Косован Г. В. – канд. техн. наук, асистент кафедри радіотехніки та інформаційної безпеки Чернівецького національного університету імені Юрія Федьковича, Чернівці, Україна.

Крояло П. М. – аспірант кафедри радіотехніки та інформаційної безпеки Чернівецького національного університету імені Юрія Федьковича, Чернівці, Україна.

АНОТАЦІЯ

Актуальність. Розглянуто задачу генерування псевдовипадкових послідовностей (ПВП) бітів із застосуванням правил нечіткої логіки та двовимірних хаотичних систем. Об'єктом дослідження є генератори псевдовипадкових бітових послідовностей побудовані із застосуванням двовимірних хаотичних систем. Мета роботи – розроблення та реалізація генераторів ПВП бітів на основі правил нечіткої логіки та двовимірних хаотичних систем та оцінка статистичних характеристик сформованих послідовностей за допомогою статистичних тестів NIST.

Метод. Запропоновано спосіб генерування псевдовипадкових послідовностей бітів, що дозволяє сформувати бітові послідовності із характеристиками, що задовольняють вимогам захищених систем зв'язку та криптографічного захисту інформації на основі правил нечіткої логіки та двовимірних хаотичних систем. В процесі дослідження роботи генераторів побудовано гістограми розподілу вихідних значень, що дозволяє чітко встановити, чи весь діапазон вихідних значень двовимірної системи може бути використаний для генерування ПВП бітів чи тільки його частина. Також проведено дослідження статистичних характеристик генерованих послідовностей за допомогою набору статистичних тестів.

Результати. Послідовності бітів сформовані із застосуванням правил нечіткої логіки та двовимірних хаотичних систем можуть бути використані для передачі інформації в захищених системах зв'язку.

Висновки. Проведені експерименти підтвердили здатність запропонованих генераторів генерувати бітові послідовності із хорошими статистичними характеристиками, що і дозволяє їх рекомендувати для використання на практиці при вирішенні задач криптографічного захисту інформації та захищеної передачі інформації по відкритих каналах зв'язку. Перспективи подальших досліджень можуть полягати в створенні криптографічних методів захисту інформації на основі запропонованих генераторів ПВП бітів, реалізації захищених систем зв'язку.

КЛЮЧОВІ СЛОВА: генератор, хаос, багатовимірна система, псевдовипадкова послідовність, нечітка логіка, статистичний тест.

УДК 004.056.55, 004.942

ИССЛЕДОВАНИЕ СВОЙСТВ ГЕНЕРАТОРОВ ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПОСТРОЕННЫХ С ИСПОЛЬЗОВАНИЕМ НЕЧЕТКОЙ ЛОГИКИ И ДВУМЕРНЫХ ХАОТИЧЕСКИХ СИСТЕМ

Кушнир Н. Я. – канд. физ.-мат. наук, доцент, доцент кафедры радиотехники и информационной безопасности. Черновицкого национального университета имени Юрия Федьковича, Черновцы, Украина.

Косован Г. В. – канд. техн. наук, ассистент кафедры радиотехники и информационной безопасности Черновецкого национального университета имени Юрия Федьковича, Черновцы, Украина.

Крояло П. М. – аспирант кафедры радиотехники и информационной безопасности Черновицкого национального университета имени Юрия Федьковича, Черновцы, Украина.

DOI 10.15588/1607-3274-2022-1-5

АННОТАЦИЯ

Актуальность. Рассмотрена задача генерирования псевдослучайных последовательностей битов (ПСП) с применением правил нечеткой логики и двумерных хаотических систем. Объектом исследования являются генераторы псевдослучайных битовых последовательностей построенных с применением двумерных хаотических систем. Цель работы – разработка и реализация генераторов ПСП бит на основе правил нечеткой логики и двумерных хаотических систем и оценка сформированных последовательностей с помощью статистических тестов NIST.

Метод. Предложен способ генерирования псевдослучайных последовательностей битов, позволяет сформировать битные последовательности с характеристиками, удовлетворяющими требованиям защищенных систем связи и криптографической защиты информации на основе правил нечеткой логики и двумерных хаотических систем. В процессе исследования работы генераторов построены гистограммы распределения выходных значений, позволяет четко установить, весь диапазон выходных значений двумерной системы может быть использован для генерирования ПСП битов или только его часть. Также проведено исследование статистических характеристик генерируемых последовательностей с помощью набора стать-соприкасающихся тестов.

Результаты. Последовательности битов сформированы с применением правил нечеткой логики и двумерных хаотических систем могут быть использована для передачи информации в защищенных системах связи.

Выводы. Проведенные эксперименты подтвердили способность предложенных генераторов генерировать битные последовательности с хорошими статистическими характеристиками, что и позволяет их рекомендовать для использования на практике при решении задач криптографической защиты информации и защищенной передачи информации по открытым каналам связи. Перспективы дальнейших исследований могут заключаться в создании криптографических методов защиты информации на основе предложенных генераторов ПСП битов, реализации защищенных систем связи.

КЛЮЧЕВЫЕ СЛОВА: генератор, хаос, многомерная система, псевдослучайная последовательность, нечеткая логика, статистический тест.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Kocarev L. Chaos-based cryptography: A brief overview / L. Kocarev // IEEE Circuits and Systems Maganize. – 2001. – Vol. 1. – P. 6–21. DOI:10.1109/7384.963463.
2. Features of creating based on chaos pseudo-random sequences. Modern Problems of Radio Enginiring, Telecommunica-tions, and Computer Sciens / [A. Semenko, N. Kushnir, N. Bokla, Hr. Kosovan] // XIIth International Conference TCSET' February 20–24 2018: proceedings. – Lviv-Slavsco, Ukraine. 2018. – P. 338–342. DOI:10.20535/2411-2976.22018.
3. Mira C. Chaotic dynamics in two-dimensional noninvertible maps / C. Mira and all // World Scientific Series on Nonlinear Science, Series A. – 1996. – Vol. 20. – P. 185–337. https://doi.org/10.1142/2252.
4. Hénaff S. Dynamical Analysis of a new statistically highly performant deterministic function for chaotic signals generation / S. Hénaff, I. Taralova, R. Lozi // International Conf. on Physics and Control (PhysCon): proceedings. – Catania, Sicily, September 2009. – P. 10. HAL Id: hal-00623064.
5. Strogatz S. H. Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering / S. H. Strogatz. – CRC Press: Boca Raton, FL, USA, 2018. – P. 532. ISBN 9780813349107.
6. New Pseudorandom Bit Generator Based on Mixing Three-Dimensional Chen Chaotic System with a Chaotic Tactics / [X. Huang, L. Liu, X. Li et al.] // Complexity. – 2019. – № 44. – P. 1–9. https://doi.org/10.1155/2019/6567198.
7. A pseudorandom number generator based on piecewise lo-gistic map / [Y. Wang, Z. Liu, J. Ma, H. He] // Nonlinear Dyn. – 2016. – № 83. – P. 2373–2391. https://doi.org/10.1007/s11071-015-2488-0.
8. A novel pseudorandom number generator based on pseudorandomly enhanced logistic map / [M. Murillo-Escobar, C. Cruz-Hernández, L. Cardoza-Avendaño, R. Méndez-Ramírez] // Nonlinear Dyn. – 2017. – № 87. – P. 407–425. https://doi.org/10.1007/s11071-016-3051-3.
9. Zimmermann H. J. Fuzzy Set Theory – And Its Applications / H. J. Zimmermann // Springer Science & Business Media : Berlin, Germany. – 2011. – Vol. 21. – 525 p. DOI: 10.1007/978-94-015-8702-0.
10. Chakraverty S. Concepts of Soft Computing: Fuzzy and ANN with Programming / S. Chakraverty, D. M. Sahoo, N. R. Mahato // Springer : Berlin/Heidelberg, Germany, 2019. – 198 p. DOI 10.1007/978-981-13-7430-2.
11. Hanss M. Applied Fuzzy Arithmetic: An Introduction with Engineering Applications / M. Hanss. – Springer : Berlin/Heidelberg, Germany, 2005. – 270 p. DOI: 10.1007/b138914.
12. Li Z. On Fuzzy Logic and Chaos Theory: from an Engineer-ing Perspective. In Fuzzy Logic / Z. Li, X. Zhang. – A Spec-trum of Theoretical & Practical Issues; Springer : Ber-lin/Heidelberg, Germany, 2007. – P. 79–97. ISSN: 1434-9922.
13. Porto M. A fuzzy approach for modeling chaotic dynamics with assigned properties / M. Porto, P. Amato // Ninth IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2000: proceedings. – San Antonio, TX, USA, 7–10 May 2000. – Vol. 1. – P. 435–440. DOI: 10.1109/FUZZY.2000.838699.
14. Stefanini L. Simulation of fuzzy dynamical systems using the LU-representation of fuzzy numbers / L. Stefanini, L. Sorini, M. L. Guerra // Chaos Solitons Fractals 2006. – № 29. –P. 638–652. https://doi.org/10.1016/j.chaos.2005.08.096.
15. Patidar V. A pseudo random bit generator based on chaotic logistic map and its statistical testing / V.Patidar, K. K. Sud, N. K. Pareek // Informatica. – 2009. – № 33. – P. 441–452.
16. Stojanovski T. Chaos-based random number generators-part I: analysis [cryptography] / T. Stojanovski, L. Kocarev // IEEE Trans. Circuits Syst. I Fundam. Theory Appl. – 2001. – № 48. – P. 281–288. DOI: 10.1109/81.915385.
17. Pseudo-random number generator based on mixing of three chaotic maps / [M. François, T. Grosges, D. Barchiesi, R. Erra] // Commun. Nonlinear Sci. Numer. Simul. – № 19. – P. 887–895. https://doi.org/10.1007/978-3-319-06089-7_16.
18. A statistical Test Suite for Random and Pseudorandom Number Generators for cryptographic Applications / [Rukhin A., Soto J., Nechvatal J., Smid M. et al.] // Technical Report; Booz-Allen and Hamilton Inc.: Mclean, VA, USA. – 2001. DOI: 10.3390/sym12081202.
19. Alvarez G. Some basic cryptographic requirements for chaos-based cryptosystems / Alvarez G., Li S. // Int. J. Bifurc. Chaos. – 2006. – № 16. – P. 2129–2151. DOI: 10.1142/S0218127406015970.
20. Encryption of the Images on the Basis of Two Chaotic Sys-tems with the use of Fuzzy Logic / [M. Kushnir, Hr. Kosovan, P. Kroialo, A. Komarnytskyy] // 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET 2020: proceedings. – Lviv-Slavsco, Ukraine. February 25–29, 2020. – P. 610–613. DOI: 10.1109/TCSET49122.2020.235504.
21. Modification of the Logistic Map Using Fuzzy Numbers with Application to Pseudorandom Number Generation and Image Encryption / Moysis L., Volos Ch., Jafari S. et al. // Entropy. – 2020. – Vol. 22. – 474 p. DOI:10.3390/e22040474.
22. A New Parallel Fuzzy Multi Modular Chaotic Logistic Map for Image Encryption / Gad M., Hagras E., Soliman H. et al. // The International Arab Journal of Information Technology. – March 2021. – Vol. 18, No. 2. – P. 227–236. https://doi.org/10.34028/iajit/18/2/12.

DOI 10.15588/1607-3274-2022-1-5

UDC 004.93

SOLVING POISSON EQUATION WITH CONVOLUTIONAL NEURAL NETWORKS

Kuzmych V. A. – PhD, Student of Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.

Novotarskyi M. A. – Dr. Sc., Professor of Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.

Nesterenko O. B. – PhD, Head of the Department of Applied Physics and Higher Mathematics, Kyiv National University of Technologies and Design, Kyiv, Ukraine.

ABSTRACT

Context. The Poisson equation is the one of fundamental differential equations, which used to simulate complex physical processes, such as fluid motion, heat transfer problems, electrodynamics, etc. Existing methods for solving boundary value problems based on the Poisson equation require an increase in computational time to achieve high accuracy. The proposed method allows solving the boundary value problem with significant acceleration under the condition of acceptable loss of accuracy.

Objective. The aim of our work is to develop artificial neural network architecture for solving a boundary value problem based on the Poisson equation with arbitrary Dirichlet and Neumann boundary conditions.

Method. The method of solving boundary value problems based on the Poisson equation using convolutional neural network is proposed. The network architecture, structure of input and output data are developed. In addition, the method of training dataset generation is described.

Results. The performance of the developed artificial neural network is compared with the performance of the numerical finite difference method for solving the boundary value problem. The results showed an acceleration of the computational speed in x10–700 times depending on the number of sampling nodes.

Conclusions. The proposed method significantly accelerated speed of solving a boundary value problem based on the Poisson equation in comparison with the numerical method. In addition, the developed approach to the design of neural network architecture allows to improve the proposed method to achieve higher accuracy in modeling the process of pressure distribution in areas of arbitrary size.

KEYWORDS: machine learning, Poisson equation, convolutional neural network.

ABBREVIATIONS

CFD is a computational fluid dynamics;
ANN is an artificial neural network;
FPN is a Feature Pyramid Network;
MSE is mean squared error;
GPU is a graphics-processing unit;
PDE is a partial differential equation;
RHS is a right hand side;
BC is a boundary condition.

NOMENCLATURE

Δ is the Laplace operator;

$F(x,y)$ is the right-hand side function;

$u(x,y)$ is the unknown function;

$u_{i,j}$ is the discrete value of the unknown function at the node with (i,j) coordinates;

h is an area sampling step;

n,m are the number of steps in x and y coordinates, respectively;

φ_i is the Dirichlet boundary condition.

INTRODUCTION

Traditionally, we use CFD modeling to determine the distribution of pressure and other parameters of the movement of liquids or gases. Finite Difference Method, Finite Element Method and Finite Volume Method are key techniques for solving aerodynamic problems, weather forecasting, life sciences and many other fields.

However, all these methods have a number of significant disadvantages. The main common disadvantage of CFD methods is a significant increase in the amount of computation and memory used with increasing number of sampling nodes or domain size.

When using the finite differences method, the solution of such a system can be simplified by using the fact that in this case we obtain a three-diagonal matrix. This structure of the matrix allows the use of parallel calculations, which significantly reduces the time to solve the problem. However, this approach makes sense to apply in the case of an area with simple geometry, because the finite difference method may lose convergence.

These factors interfere the widespread use of numerical methods in real-time applications and stimulate active research into alternative methods for solving PDEs to overcome these limitations.

The use of ANN has great prospects for solving these problems. This opportunity appeared through significant progress in the field of deep machine learning. The use of ANN of various types to solve boundary value problems is developing rapidly. This progress is based on a real opportunity to overcome those objective limitations that numerical methods have.

One of the fields of study, where deep learning algorithms can be applied is a biomedical engineering. Accuracy of modelling of biological objects has less crucial role, than speed of modelling. Moreover, GPU-accelerated neural networks can be a new efficient solution for many biomedical problems.

DOI 10.15588/1607-3274-2022-1-6

Example of such problem is the mathematical models of anastomoses of human stomach. Determination of the pressure field in the zone, which can be described by the Poisson equation, of the reconstruction-recovery operation in the real time, doesn't require extremely high precision. That's why we consider application of artificial neural networks as alternative to numerical methods to solve mentioned problem.

The object of study is the process of computational hydrodynamics.

The subject of study is the methods and means of using ANN to solve the Poisson equation.

The aim of this work is to develop and train an artificial neural network for modeling pressure field changes in areas with complex geometry and boundary conditions, which allows to reduce the simulation time with acceptable accuracy.

1 PROBLEM STATEMENT

Poisson equation is an elliptic partial differential equation. The solution of this equation is presented as a boundary value problem in a rectangular domain Ω with parameters $(x, y) \in [0, n] \times [0, m]$. This problem can be mathematically represented by next formulas:

$$\begin{cases} \Delta u(x, y) = F(x, y), \\ u(x, 0) = \varphi_1(x), \\ u(0, y) = \varphi_2(y), \\ u(x, m) = \varphi_3(x), \\ u(n, y) = \varphi_4(y). \end{cases} \quad (1)$$

To construct a difference scheme for the boundary value problem (1), we introduce a uniform grid: $\omega_h = \{x_i = ih, y_j = jh, i = \overline{1, n}, j = \overline{jm}\}$, where h=1.

After discretization of the boundary value problem (1), using the symbolic notation of difference operators, we obtain the difference boundary value problem on the ω_h grid:

$$\begin{cases} -(\Delta_h) u_{i,j} = f_{i,j}, \\ u_{i,0} = \varphi_1(i), \\ u_{0,j} = \varphi_2(j), \\ u_{i,m} = \varphi_3(i), \\ u_{n,j} = \varphi_4(j). \end{cases} \quad (2)$$

We use the second central difference to approximate the Laplace operator at an arbitrary interior cell with coordinates (x_i, y_j) and perform calculations on a five-point template. Thus, we obtain an approximation of the Poisson equation for interior cells:

$$-(\Delta_h u)_{i,j} = \frac{4u_{i,j} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1}}{h^2} = f_{i,j} \quad (3)$$

Since we use a step size equal to one, equation (3) has a simplified form:

$$f_{i,j} = 4u_{i,j} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1}. \quad (4)$$

In reality, we do not use the whole domain Ω to get useful results, but only a certain part of it. The object in the domain Ω has a complex shape bounded by obstacles. To represent difference between regions, the $(m \times n)$-matrix $dist$ introduced by its elements:

$$dist_{ij} = \begin{cases} 1, (i, j) - \text{node of the obstacle region}, \\ 0, (i, j) - \text{node of the object region}. \end{cases} \quad (5)$$

For all (i, j)-cells with $dist_{ij} = 1$ we use $f_{i,j} = 0$. Otherwise, function $f_{i,j}$ defined in tabular form. The values of the $f_{i,j}$ function can vary in the range [–1, 1]. Example of domain, divided into obstacle and object, presented in Fig. 1. Part of such domain presented more detail in Fig. 2.

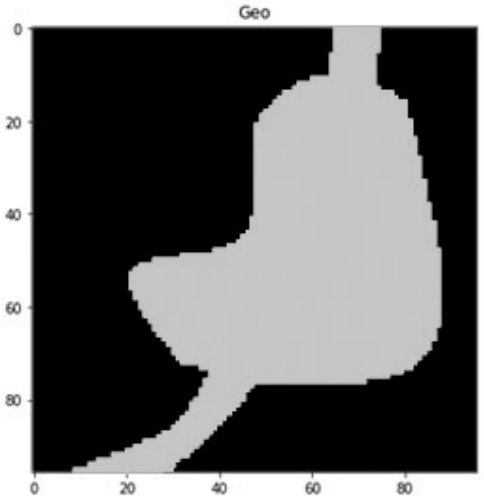


Figure 1 – The black region is an obstacle and the gray region is an object

We carry out calculations in accordance with equation (4) only for those (*i, j*)-cells for which the following conditions are satisfied:

$$\begin{cases} dist_{i,j+1} = 0, \\ dist_{i,j-1} = 0, \\ dist_{i+1,j} = 0, \\ dist_{i-1,j} = 0. \end{cases} (1 < i < n, 1 < j < m).$$

DOI 10.15588/1607-3274-2022-1-6

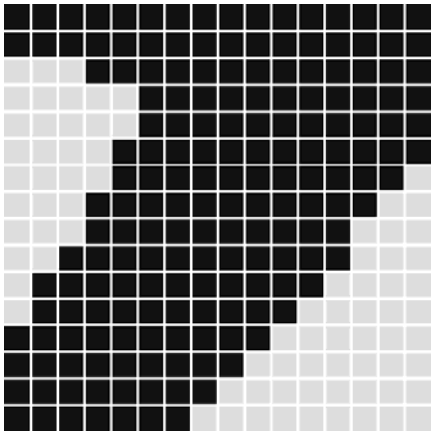


Figure 2 – The black squares are obstacle cells and the gray squares are object cells

The remaining cells belonging to the object may have 1, 2 or 3 neighbors located in the obstacle region. There may be various combinations of obstacle and object neighbor cells. All of them can be obtained by rotating and/or mirroring the configurations shown in Fig.3.

We use different approximation schemes for each of the following configurations:

– $4u_{i,j} - u_{i-1,j} - u_{i+1,j} - 2u_{i,j-1} = f_{i,j}$ for the cell with one obstacle (Fig. 3a),

– $4u_{i,j} - u_{i,j+1} - u_{i,j-1} = f_{i,j}$ for two opposite obstacles (Fig. 3b),

– $4u_{i,j} - 2u_{i,j-1} - 2u_{i+1,j} = f_{i,j}$ for two non-opposite obstacles (Fig. 3c),

– $4u_{i,j} - 2u_{i,j+1} = f_{i,j}$ for three obstacles (Fig. 3d),

We assume that $f_{i,j} = 0$ if i and j are outside the domain.

To solve this boundary value problem, we transform (2) into a system of linear algebraic equations and calculate the unknown value for each cell.

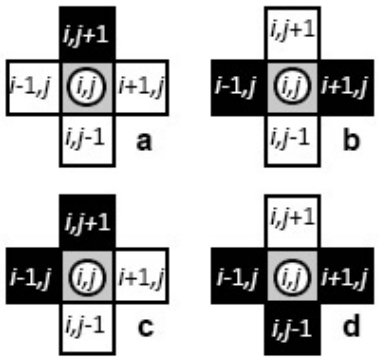


Figure 3 – Basic combinations of obstacle and object neighbor cells

2 REVIEW OF THE LITERATURE

The dynamics of fluid motion is an extremely important task that underlies research and engineering solutions in various fields of science and technology. CFD is widely used in many fields of study and industries [1]. At the same time, there is a problem associated with the fact that all numerical methods have a number of significant disadvantages [2].These disadvantages can be partially overcome by parallelizing the finite difference method [3]. However, to achieve convergence of this method, we must only use simple domain geometry or adaptive meshes [4].

Thanks to the successful development of machine learning and deep neural networks, it has become possible to make a breakthrough in the field of modeling physical processes [5]. The first investigations into solving PDEs using machine learning was carried out in the early 90s. Dissanayake and Phan-Thien [6] proposed multilayer perceptron to solve non-linear problems, such as Poisson equation and thermal conduction with non-linear heat generation. The main limitation of this method is a relatively small size of modeled domain [7]. Lee and Kang [8] introduced the general concept of developing neural algorithms for solving differential equations. Main limitations of studies of mentioned period are specific range of modelled RHS functions, a narrow set of BC and small domain sizes.

In the beginning of 2000s, with the improvement of computational efficiency, much more studies were published, which presented more complicated and robust models. In [9] the authors investigated Multilayer perceptrons to predict the orthogonal decomposition of the 2D Navier-Stokes equation and the 1D Kuramoto-Sivashinsky equation. A more in-depth and complex review of methods and techniques was conducted by Yadav and Kumar [10].

Xiao et al. [11] and Tompson et al. [12] were one of the first to research the use of convolutional neural networks to solve the Poisson equation. Both works proposed similar approaches to solve a boundary value problem based on the Poisson equation with a given RHS function. It is reported that in both cases there are problems with the accuracy of the results with reduced simulation time.

3 MATERIALS AND METHODS

The structure of the developed ANN was adapted and modified from the FPN architecture [13]. The presented version of the Poisson solution-oriented ANN allowed reducing the number of trained parameters from 23534592 in the basic network to 337447.

The presented ANN includes two main parts, called “bottom-up pathway” and “top-down pathway”. The general structure of ANN is shown in Fig. 4.

The input data for the “bottom-up pathway” is presented as a tensor of size 96×96×2. We obtain such data by combining two 96×96 matrices, where the first matrix is the RHS of the Poisson equation and the second matrix encodes the geometric space. The values of the elements of the second matrix correspond to the following rules: if the grid cell is located in the obstacle area, then the corresponding matrix element is zero, otherwise the value of the matrix element is equal to the distance to the nearest

DOI 10.15588/1607-3274-2022-1-6

obstacle cell divided by 96√2. Example of geometry and respective encoding is shown in Fig. 5.

The "bottom-up pathway" part consists of 4 blocks. There are three modified "residual" blocks and one "pre_conv" block in it. The structure of the "pre_conv" block is shown in Fig. 6. It contains three layers, namely 2D convolution layer, batch normalization layer and relu activation layer. We included this block in the ANN structure to increase the number of channels from 1 to 64 and pass the tensor forward to the "residual" blocks.

Each subsequent "residual" block halves the tensor obtained from the previous block and transmits the result to the corresponding upsampling layer, which is located in the "top-down pathway" part. The proposed structure of a convolutional neural network uses the main property of neural networks of this type, which combines the characteristics of the fluid at different scales. Thus, we combine each cell with the rest of the considered domain. The data outputs of each "residual" block and the "pre_conv" block are transmitted to a 2D convolution layer with a core size of 1 × 1 to reduce the number of channels.

Next, convolution outputs pass to upsampling layers, which form "top-down pathway", where size of first "residual" block output multiplies by 2, second output by 4, and third output by 8. After this operation, all outputs combine into single tensor by concatenating them along last axis. Finally, this tensor moves to a two-dimensional block "output_conv", which consists of 2 layers of convolution: the first layer consists of 7 filters and core size 3, and the second contains 1 filter, core size 3 and the function of activating the hyperbolic tangent in the range from –1 to 1. Output tensor of the model is 2D array, with shape 96×96. To obtain solution of Poisson equation with input RHS, values of the output array must be rescaled into the original training data distribution.

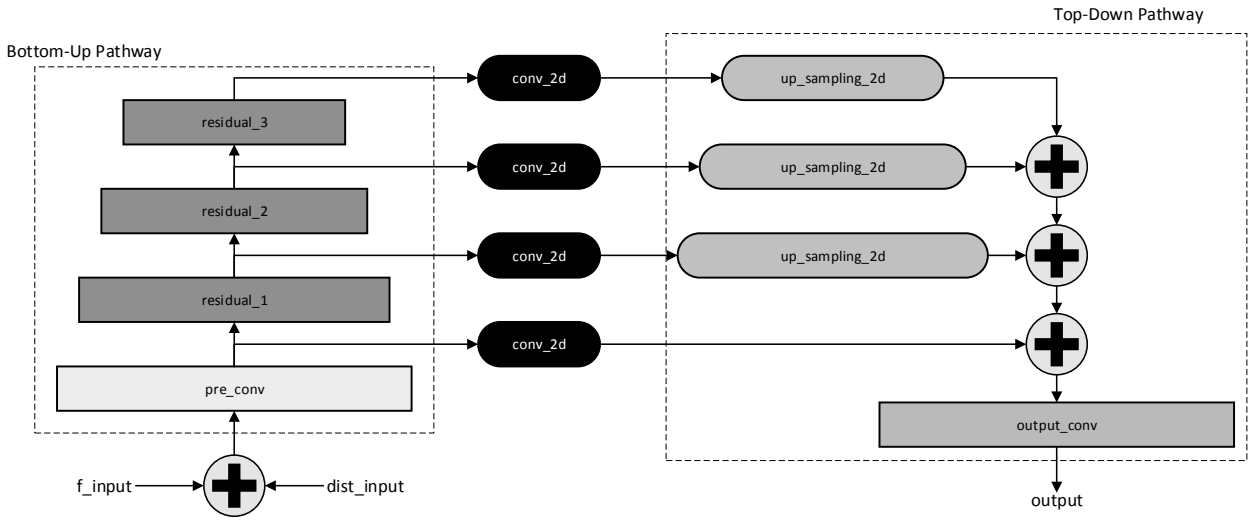


Figure 4 – General structure of the ANN

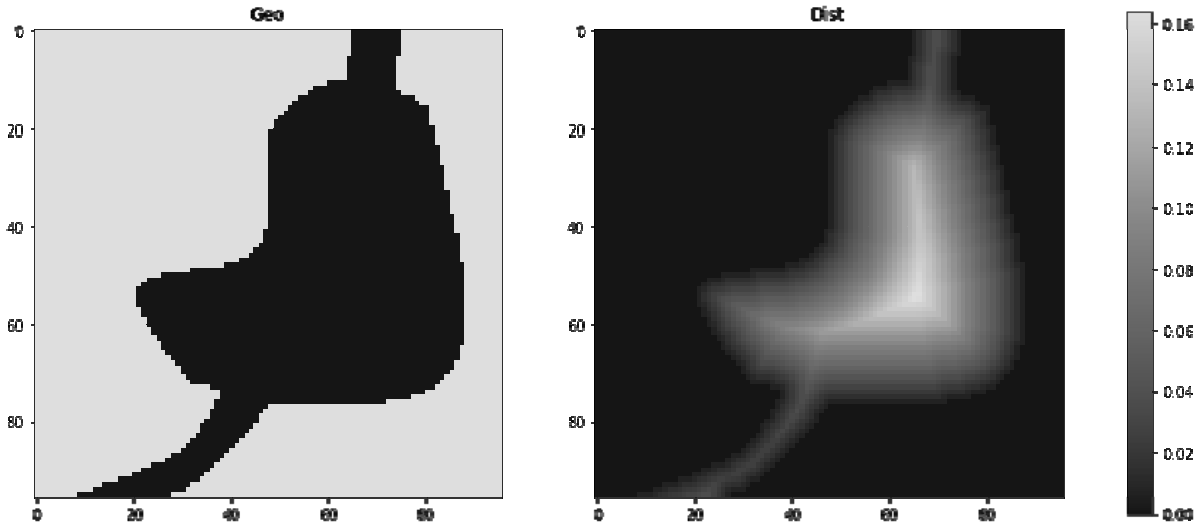


Figure 5 – The left shows the geometry of the zone where the light shade represents the obstacle zone; the light shades in the right figure correspond to the values of the elements of the distance matrix to the nearest obstacle

DOI 10.15588/1607-3274-2022-1-6

The training data set contains 40,000 pairs of samples, where the features are represented by 2 matrices of size 96 × 96. One matrix contains a tabular representation of RHS, and the other specifies the geometry of the space. The target matrix is a solution of the Poisson equation of 96 × 96 size. The solution is computed using the algebraic multigrid Python PyAMG package [14].

We prepared 140 geometries, and for each geometry, we generated 250 RHS matrices and corresponding solutions. Thus, 35 000 samples were processed.

Examples of such a pairs presented in fig. 8. Additionally, we generated 5000 samples without any geometries. An example of such a pair we can see in Fig. 7. In this case, all values of the distance matrix were equal to 1. RHS matrices were created by generating 8 × 8, 12 × 12 and 16 × 16 low-resolution grids with random values from –1 to 1, and then increasing the generated grids to the target size of 96 × 96 by cubic interpolation.

Each pair of samples contains an array with the solution of the equation, i.e. the target value. However, those values constrained in an interval, that highly bigger, that [–1, 1]. To provide stability of network training, all values of generated solution were normalized to distribution with 0 mean and 1 standard deviation. After normalization, those values were rescaled to range [–1, 1]. Data normalization and scaling was performed by scikit-learn library [15]. Those values are final target variable of training process.

The model was implemented using TensorFlow 2.4.1, with Adam weight optimizer [16], with following parameters: learning_rate=0.001, beta_1 = 0.9, beta_2 = 0.999, epsilon = 1e–7. We take mean square error as a loss function during training.

Training was conducted on GPU MSI GeForce GTX 1660 Super Ventus OC 6GB GDDR6, during 300 epochs and with batch size of 32.

4 EXPERIMENTS AND RESULTS

The described ANN is used in modeling the distribution of pressure in the human stomach. We modeled the stomach in 3 states-normal state and 2 different types of anastomosis (Fig. 9).

MSE in the first case is 0.000185, in the second – 0.000161, in the third – 0.000344.

Results demonstrated in Fig. 10–12.

We obtained an increase in the simulation speed compared to the numerical method. The PyAmg package was used to implement the numerical method. We measured the time of solving the boundary value problem on 1, 10, 50, 100, 200, 500 and 1000 samples. Due to the ability of ANN to process many samples simultaneously using a graphics processor, the highest increase in acceleration speed was achieved in 500 samples (Table 1 and Fig. 13).

5 DISCUSSION

Figures 10 and 11 show that the trained neural network selects the same regions with extreme pressure values as the numerical method chosen as the ground truth. Despite the ability to distinguish areas with high or low values, the trained ANN needs to be improved to make better predictions in the direction of smoothly varying values.

The experiments showed an increase in the simulation speed compared to the chosen numerical method. This fact can also be explained in particular by conducting experiments using GPU, which accelerate the matrix operations that are basic for CNN. Moreover, used GPU – nVidia GTX 1660 Super isn't the fastest GPU nowadays. Using newer and powerful GPU will achieve even bigger gain in computational speed. In respect of this fact and relatively small accuracy loss, in comparison with numerical method, the developed method can be applied in fields, where speed of modelling is more crucial, than accuracy, like some parts of biomedical engineering etc.

Further research will be aimed on improving neural network architecture to make predictions smoother. In addition, we believe that the main disadvantage of this result is the fixed size of the domain of 96 × 96. Therefore, we will focus on developing methods of deep learning paradigms that will allow us to work with arbitrary domain sizes. This approach involves improving the generation of training data sets, including the choice of a numerical method for solving a boundary value problem that directly affects the accuracy of the neural network.

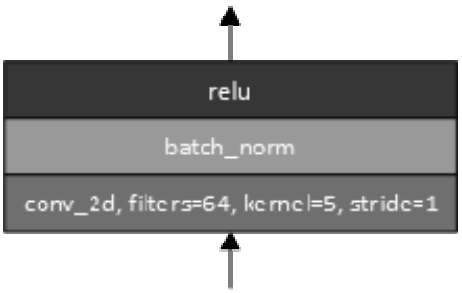


Figure 6 – "Pre_conv" block

DOI 10.15588/1607-3274-2022-1-6

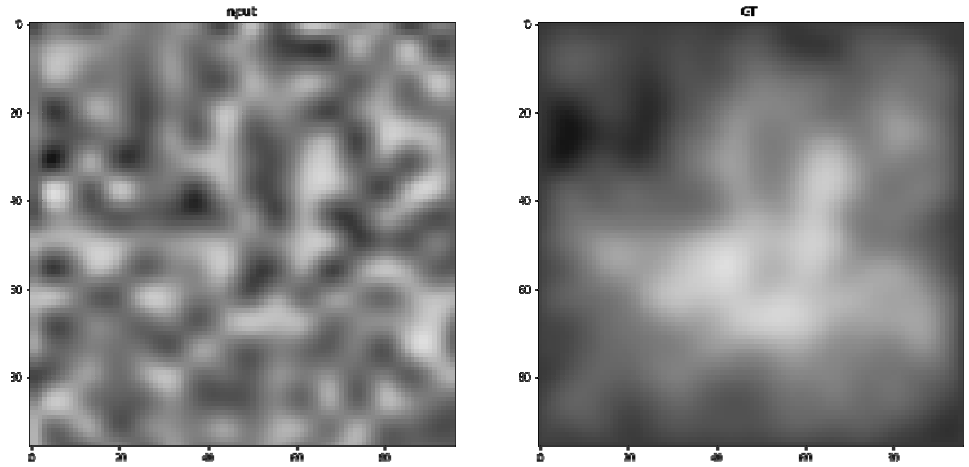


Figure 7 – The graphical representation of the RHS function is presented on the left, the corresponding solution is shown on the right

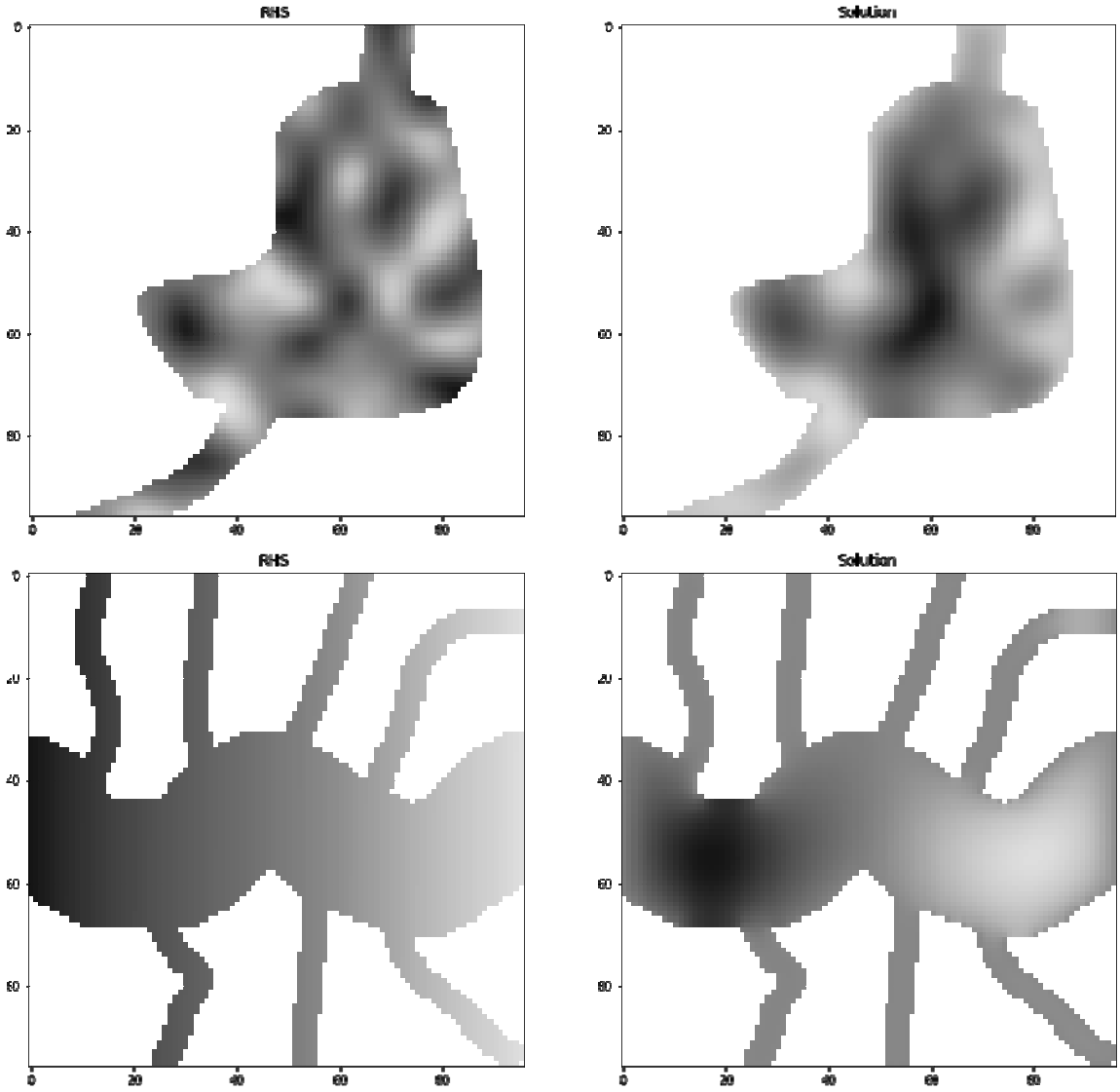


Figure 8 – RHSs for the equation are shown on the left, the corresponding solutions are shown on the right; white indicates obstacles

DOI 10.15588/1607-3274-2022-1-6

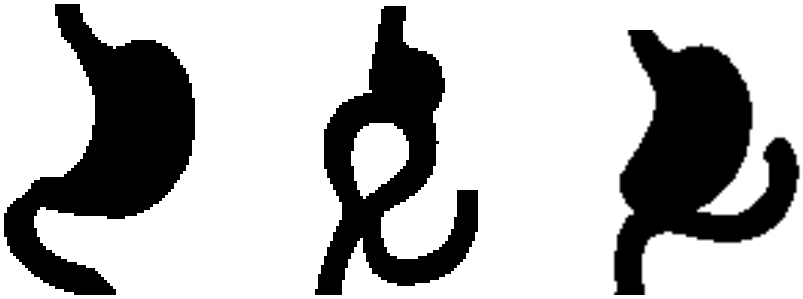


Figure 9 – Left –normal stomach, center and right-anastomosis

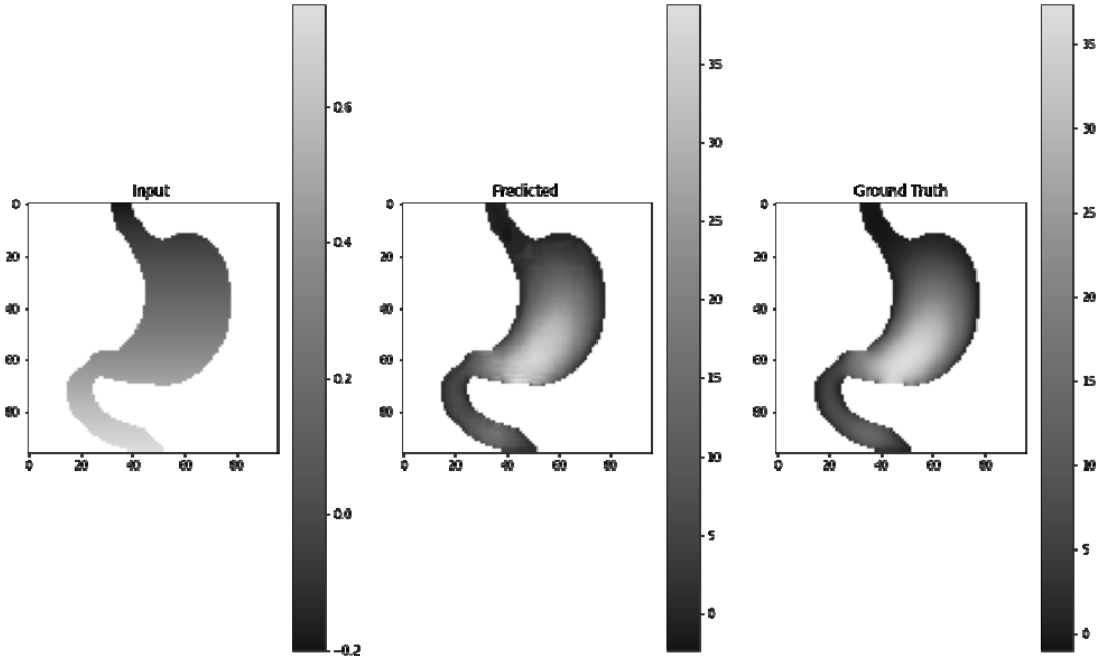


Figure 10 – Normal stomach; left – RHS of equation, center – network prediction, right – ground truth; white color denotes obstacles

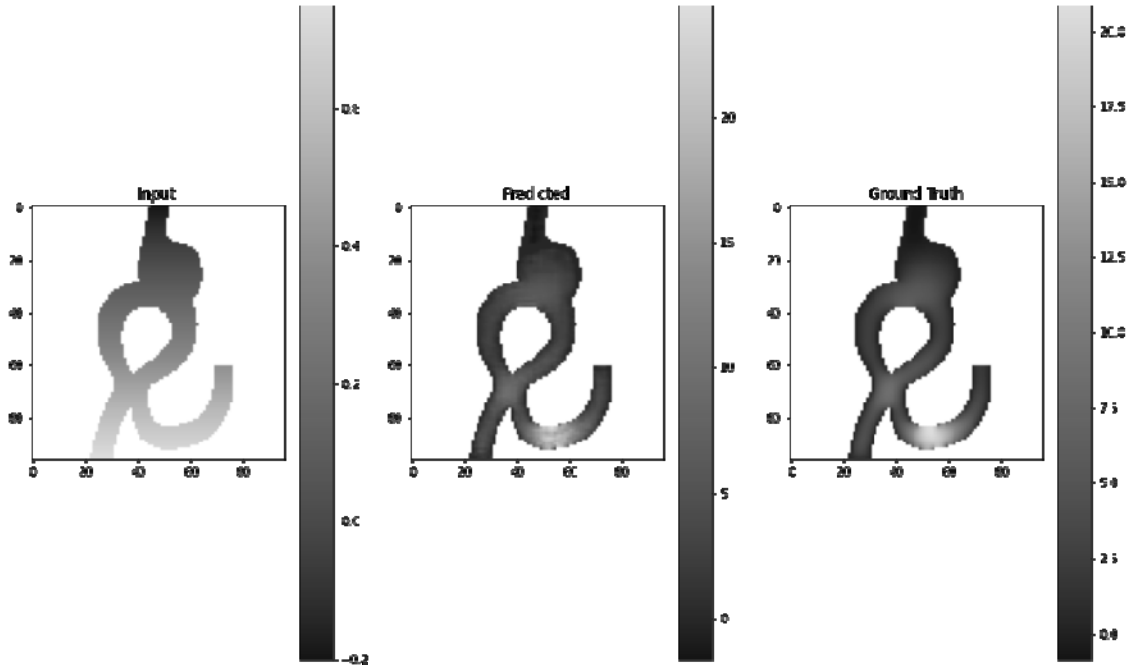


Figure 11 – Anastomosis; on the left – the right equation, the center – the network forecast, on the right – the basic truth; white indicates obstacles

DOI 10.15588/1607-3274-2022-1-6

Table 1 – Comparison of ANN and numerical method speed performance

Number of samples	Total time – ANN, ms	Time per sample – ANN, ms	Total time – numerical method, ms	Time per sample – numerical method, ms
1	23.80	23.80	237.00	237.00
10	29.00	2.90	7460.00	746.00
50	55.80	1.12	31400.00	628.00
100	91.20	0.91	52600.00	526.00
200	159.00	0.80	99000.00	495.00
500	334.00	0.67	261000.00	522.00
1000	725.00	0.73	536000.00	536.00

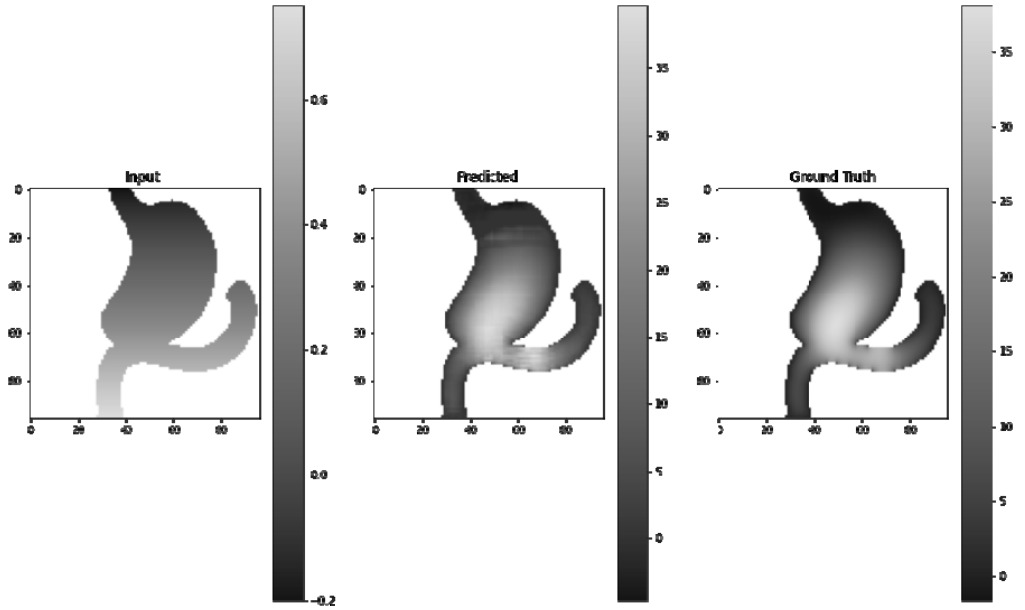


Figure 12 – Anastomosis; on the left – the right equation, the center – the network forecast, on the right – the basic truth; white indicates obstacles

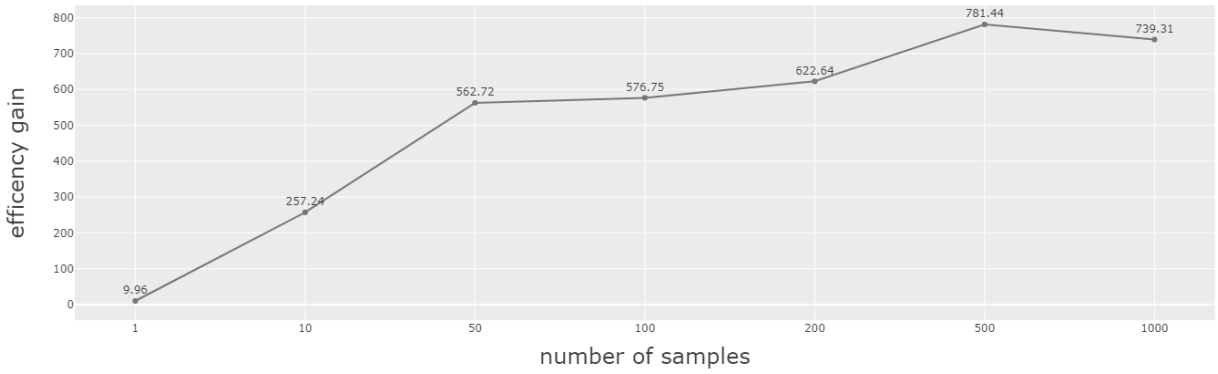


Figure 13 – Dependence of efficiency gain on the number of samples

CONCLUSIONS

This paper has investigated application of convolutional neural networks in solving boundary value problem based on the Poisson equation. We propose novel approach to solve Poisson equation with Dirichlet and Neumann boundary conditions, in domain with fixed size – 96×96.

Experiments with the developed method shown big gain of computational speed, in comparison with numerical method.

Scientific novelty is represented by the method of solving a boundary value problem based on the Poisson equation, which allowed to significantly accelerate the speed of its solution in comparison with the numerical

DOI 10.15588/1607-3274-2022-1-6

method under the condition of a given acceptable accuracy.

The practical value of this achievement and insignificant loss in prediction accuracy is that convolutional neural network based technique can be applied in different field of engineering and science, where modelling speed matters.

Future work will focus on improvement of generalizability of CNN architecture, in order to handle various domain shapes and sizes. In addition, improving the speed of simulations using GPU for the tasks under consideration has not been widely studied. Therefore, continued research in this area is extremely important.

ACKNOWLEDGEMENTS

The results of these studies were used in scientific research projects "Artificial Intelligence Platform for Distant Computer-Aided Detection and Computer-Aided Diagnosis of Human Diseases" (state registration number 0120U105119) and "Application of mathematical methods in the problems of natural and technical sciences" (state registration number 0120U100979).

REFERENCES

1. Tu J. Computational Fluid Dynamics: A Practical Approach. Oxford: Butterworth-Heinemann, 2018, 495 p.
2. Wu C. Y., Ferng Y. M., Ciieng C. C., Liu C. C. Investigating the advantages and disadvantages of realistic approach and porous approach for closely packed pebbles in CFD simulation, *Nuclear Engineering and Design*, 2010, Vol. 240, pp. 1151–1159.
3. Simon H. D., Gropp W., Lusk E. Parallel Computational Fluid Dynamics: Implementations and Results (Scientific and Engineering Computation). Cambridge, The MIT Press, 1992, 362 p.
4. Runnels B., Agrawal V., Zhang W., Almgren A. Massively parallel finite difference elasticity using block-structured adaptive mesh refinement with a geometric multigrid solver, ArXiv e-prints, 2020, https://arxiv.org/abs/2001.04789v2, DOI: https://doi.org/10.1016/j.jcp.2020.110065
5. Raghu M., Schmidt E. A Survey of Deep Learning for Scientific Discovery, *ArXiv e-prints*, 2020, https://arxiv.org/abs/2003.11755v1.
6. Dissanayake M. W. M. G., Phan-Thien N. Neural-network-based approximations for solving partial differential equations, *Communications in Numerical Methods in Engineering,* 1994, Vol. 10, No. 3, pp. 195–201.
7. Sirignano J., Spiliopoulos K. DGM: A deep learning algorithm for solving partial differential equations, *Journal of Computational Physics*, 2018, Vol. 375, pp. 1339–1364.
8. Lee H., Kang I. S. Neural algorithm for solving differential equations, *Journal of Computational Physics*, 1990, Vol. 91, No. 1, pp. 110–131.
9. Smaoui N., Al-Enezi S. Modelling the dynamics of nonlinear partial differential equations using neural networks, *Journal of Computational and Applied Mathematics*, 2004, Vol. 170, No. 1, pp. 27–58.
10. Kumar M. and Yadav N. Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations: A survey, *Computers & Mathematics with Applications*, 2011, Vol. 62, №10, pp. 3796–3811.
11. Xiao X., Zhou Y., Wang H., and Yang X. A novel cnn-based poisson solver for fluid simulation, *IEEE Transactions on Visualization and Computer Graphics,* 2020, Vol. 26, No. 3, pp. 1454–1465. DOI: 10.1109/TVCG.2018.2873375
12. Tompson J., Schlachter K., Sprechmann P., and Perlin K. Accelerating eulerian fluid simulation with convolutional networks, *ArXiv e-prints,* 2017, https://arxiv.org/abs/1607.03597.
13. Lin T.-Y., Dollar P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection, *ArXiv e-prints*, 2017, https://arxiv.org/abs/1612.03144v2
14. Olson L. N., Schroder J. B., PyAMG: Algebraic multigrid solvers in Python v4.0.0, 2018, https://github.com/pyamg
15. Pedregosa F. et al. Scikit-learn: machine learning in Python, *Journal of Machine Learning Research*, 2011, Vol. 12, pp. 2825–2830.
16. Duchi J., Hazan E., Singer Y. Adaptive subgradient methods for online learning and stochastic optimization, *Journal of Machine Learning Research*, 2011, Vol. 12, pp. 2121–2159.

Received 16.11.2021.
Accepted 16.01.2022.

УДК 004.93

РОЗВ'ЯЗУВАННЯ РІВНЯННЯ ПУАССОНА З ЗАСТОСУВАННЯМ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Кузьмич В. А. – PhD, студент кафедри обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Новотарський М. А. – д-р техн. наук, професор кафедри обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Нестеренко О. Б. – канд. фіз.-мат. наук, завідувач кафедри прикладної фізики та вищої математики, Київський національний університет технологій та дизайну.

АНОТАЦІЯ

Актуальність. Рівняння Пуассона – це одне з фундаментальних диференціальних рівнянь, яке використовується для моделювання складних фізичних процесів, таких як рух рідини, проблеми теплообміну, електродинаміки тощо. Існуючі методи розв'язування крайових задач на основі рівняння Пуассона для досягнення високої точності, вимагають збільшення часу обчислень. Запропонований метод дозволяє розв'язувати крайову задачу зі значним прискоренням, за умови незначної втрати точності.

Мета. Метою нашої роботи є розробка архітектури штучної нейронної мережі для розв'язування крайової задачі на основі рівняння Пуассона з довільними крайовими умовами Діріхле та Неймана.

Метод. Запропоновано метод розв'язування крайових задач на основі рівняння Пуассона за допомогою згорткової нейронної мережі. Розроблено архітектуру мережі, структуру вхідних та вихідних даних. Також описано метод формування навчального набору даних.

Результати. Результати роботи розробленої нейронної мережі були порівняні з продуктивністю чисельного методу скінченних різниць для вирішення крайової задачі. Результати продемонстрували прискорення обчислювальної швидкості у x10–700 разів, в залежності від кількості вузлів дискретизації.

DOI 10.15588/1607-3274-2022-1-6

Висновки. Запропонований метод значно прискорив швидкість вирішення крайової задачі на основі рівняння Пуассона в порівнянні з чисельним методом. Також розроблений підхід до проектування архітектури нейронної мережі дозволяє вдосконалити запропонований метод для досягнення більш високої точності при моделюванні процесу розподілу тиску у областях довільного розміру.

КЛЮЧОВІ СЛОВА: машинне навчання, рівняння Пуассона, згорткова нейронна мережа.

УДК 004.93

РЕШЕНИЕ УРАВНЕНИЯ ПУАССОНА С ПРИМЕНЕНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Кузьмич В. А. – PhD студент кафедры вычислительной техники, Национальный технический университет Украины «Киевский политехнический институт имени Игоря Сикорского».

Новотарский М. А. – д-р техн. наук, профессор кафедры вычислительной техники, Национальный технический университет Украины «Киевский политехнический институт имени Игоря Сикорского».

Нестеренко О. Б. – канд. физ.-мат. наук, заведующая кафедрой прикладной физики и высшей математики, Киевский национальный университет технологий и дизайна.

АННОТАЦИЯ

Актуальность. Уравнение Пуассона – это одно из фундаментальных дифференциальных уравнений, которое используется для моделирования сложных физических процессов, таких как движение жидкости, проблемы теплообмена, электродинамики и тому подобное. Существующие методы решения краевых задач на основе уравнения Пуассона для достижения высокой точности, требуют увеличения времени вычислений. Предложенный метод позволяет решать краевую задачу со значительным ускорением при условии незначительной потери точности.

Цель. Целью нашей работы является разработка архитектуры искусственной нейронной сети для решения краевой задачи на основе уравнения Пуассона с произвольными граничными условиями Дирихле и Неймана.

Метод. Предложен метод решения краевых задач на основе уравнения Пуассона с помощью сверточной нейронной сети. Разработана архитектура сети, структура входных и выходных данных. Также описан метод формирования учебного набора данных.

Результаты. Результаты работы разработанной нейронной сети были сравнены с производительностью численного метода конечных разностей для решения краевой задачи. Результаты продемонстрировали ускорение вычислительной скорости в x10–700 раз, в зависимости от количества узлов дискретизации.

Выводы. Предлагаемый метод значительно увеличил скорость решения краевой задачи на основе уравнения Пуассона по сравнению с численным методом. Также разработанный подход к проектированию архитектуры нейронной сети позволяет улучшить предложенный метод для достижения большей точности при моделировании процесса распределения давления в областях произвольного размера.

КЛЮЧЕВЫЕ СЛОВА: машинное обучение, уравнение Пуассона, сверточная нейронная сеть.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Tu J. Computational Fluid Dynamics: A Practical Approach / J. Tu. – Oxford: Butterworth-Heinemann, 2018. – 495 p.
2. Investigating the advantages and disadvantages of realistic approach and porous approach for closely packed pebbles in CFD simulation / [Wu C. Y., Ferng Y. M., C. C. Ciieng, C. C. Liu] // Nuclear Engineering and Design. – 2010. – Vol. 240. – P. 1151–1159.
3. Simon H. D. Parallel Computational Fluid Dynamics: Implementations and Results (Scientific and Engineering Computation) / H. D. Simon, W. Gropp, E. Lusk. – Cambridge : The MIT Press, 1992. – 362 p.
4. Massively parallel finite difference elasticity using block-structured adaptive mesh refinement with a geometric multigrid solver / [B. Rennels, V. Agrawal, W. Zhang, A. Almgren] // ArXiv e-prints. – 2020. https://arxiv.org/abs/2001.04789v2, DOI:https://doi.org/10.1016/j.jcp.2020.110065
5. Raghu M. A Survey of Deep Learning for Scientific Discovery / M. Raghu, E. Schmidt // ArXiv e-prints. – 2020, https://arxiv.org/abs/2003.11755v1.
6. Dissanayake M. W. M. G. Neural-network-based approximations for solving partial differential equations / M. W. M. G. Dissanayake N. Phan-Thien // Communications in Numerical Methods in Engineering. – 1994. – Vol. 10, № 3. – P. 195–201.
7. Sirignano J. DGM: A deep learning algorithm for solving partial differential equations / J. Sirignano, K. Spiliopoulos // Journal of Computational Physics, 2018. – Vol. 375. – P. 1339–1364.
8. Lee H. Neural algorithm for solving differential equations / H. Lee, I. S. Kang // Journal of Computational Physics, 1990. – Vol. 91, №1. – P. 110–131.
9. Smaoui N. Modelling the dynamics of nonlinear partial differential equations using neural networks / N. Smaoui, S. Al-Enezi // Journal of Computational and Applied Mathematics. – 2004. – Vol. 170, №1. – P. 27–58.
10. Kumar M. Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations: A survey / M. Kumar and N. Yadav // Computers & Mathematics with Applications, 2011. – Vol. 62, № 10. – P. 3796–3811.
11. A novel cnn-based poisson solver for fluid simulation / [X. Xiao, Y. Zhou, H. Wang, and X. Yang] // IEEE Transactions on Visualization and Computer Graphics. – 2020. – Vol. 26, № 3. – P. 1454–1465. DOI: 10.1109/TVCG.2018.2873375
12. Accelerating eulerian fluid simulation with convolutional networks / [J. Tompson, K. Schlachter, P. Sprechmann, and K. Perlin] // ArXiv e-prints. – 2017. https://arxiv.org/abs/1607.03597.
13. Feature pyramid networks for object detection / [P. Dollar, R. Girshick, K. He et al] // ArXiv e-prints. – 2017, https://arxiv.org/abs/1612.03144v2
14. Olson L. N. PyAMG: Algebraic multigrid solvers in Python v4.0.0 / L. N. Olson, J. B. Schroder. – 2018. https://github.com/pyamg
15. Pedregosa F. Scikit-learn: machine learning in Python / F. Pedregosa et al. // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2825–2830.
16. Duchi J. Adaptive subgradient methods for online learning and stochastic optimization / J. Duchi, E. Hazan, Y. Singer // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2121–2159.

DOI 10.15588/1607-3274-2022-1-6

e-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2022. № 1
p-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2022. № 1

UDC 004.421

THE MODULAR EXPONENTIATION WITH PRECOMPUTATION OF REDUSED SET OF RESIDUES FOR FIXED-BASE

Prots'ko I. – Dr. Sc., Associate Professor, Department of Automated Control Systems, Lviv National Polytechnic University, Lviv, Ukraine.

Gryshchuk O. – Software Developer, LtdC "SoftServe", Lviv, Ukraine.

ABSTRACT

Context. Modular exponentiation is an important operation in many applications that requires a large number of calculations Fast computations of the modular exponentiation are extremely necessary for efficient computations in theoretical-numerical transforms, for provide high crypto capability of information data and in many other applications.

Objective – the runtime analysis of software functions for computation of modular exponentiation of the developed program that uses the precomputation of redused set of residuals for fixed-base.

Method. Modular exponentiation is implemented using of the development of the right-to-left binary exponentiation method for a fixed basis with precomputation of redused set of residuals. To efficient compute the modular exponentiation over big numbers, the property of a periodicity for the sequence of residuals of a fixed base with exponents equal to an integer power of two is used.

Results. Comparison of the runtimes of five variants of functions for computing the modular exponentiation is performed. In the algorithm with precomputation of redused set of residuals for fixed-base provide faster computation of modular exponentiation for values larger than 1K binary digits compared to the functions of modular exponentiation of the MPIR and Crypto++ libraries. The MPIR library with an integer data type with the number of binary digits from 256 to 2048 bits is used to develop an algorithm for computing the modular exponentiation.

Conclusions. In the work has been considered and analysed the developed software implementation of the computation of modular exponentiation on universal computer systems. One of the ways to implement the speedup of computing modular exponentiation is developing algorithms that can use the precomputation of redused set of residuals for fixed-base. The software implementation of modular exponentiation with increasing from 1K the number of binary digit of exponent shows an improvement of computation time with comparison with the functions of modular exponentiation of the MPIR and Crypto++ libraries.

KEYWORDS: modular exponentiation, big numbers, exponentiation algorithm, fixed-base exponentiation, residual set.

ABBREVIATIONS

GMP is a GNU Multiple Precision Arithmetic library;

ME is a modular exponentiation;

MPIR is a Multiple Precision Integers and Rationals library.

NOMENCLATURE

A is a base integer value;
b is a binary representation of the exponent x;
Base is an identifier of a base;
e_i is a part of binary representation x;
exp is an identifier of an exponent;
$ind_R A$ is an index of residue;
k is a bitlength of a value x
m is a number of the parts of binary representation x;
mod is an identifier of modulo;
N is an integer value of modulo;
P is an odd prime;
q is a positive integer;
r is a bitlength of a part of binary representation x;
r_i is a residue;
R is a primitive root;
T' is a period of the residues;
u is an offset of a period of the residues;
x is an integer value of an exponent;
x_i is an bit value of an exponent;
y is an integer value of modular exponentiation;
$\varphi(N)$ is the Euler's function.

INTRODUCTION

The task of developing an effective computational algorithm for ME for big numbers is relevant enough to solve the problems of modern asymmetric cryptography, for efficient computation of number-theoretic transforms, digital signatures and other applications [1].

The object of study is the process of analysis the developed software implementation of the computation of ME. To efficient compute the ME over large numbers the property of the periodicity of the sequence of residuals for the exponent of the fixed-basis equal to the integer power of two are used.

The subject of study is the computation of ME based on the use the bits of the binary exponent with the precomputation of redused set of residuals for fixed-base.

The purpose of the work is to increase the speed of computation of ME based of computer systems in comparison with the function of ME of the MPIR and Crypto ++ libraries.

1 PROBLEM STATEMENT

The ME and the discrete logarithm are important operations that require a large number of calculations. The problem of discrete logarithm [1] is formulated so that for known integers A, N, y find the integer x, $(A, N) = 1$; $A, N, y, x \in Z$) such that

$$x = \log A^y, (0 \le x \le N-1). \quad (1)$$

DOI 10.15588/1607-3274-2022-1-7

The number $x > 0$ is called the discrete logarithm of the number y based on A and modulo N according to formula (1).

The solution of the discrete logarithm problem can be the solution of the equation

$$A^x \bmod N = y. \quad (2)$$

That is, determining the number x, which is the solution of equation (2), we find the discrete logarithm. Thus, the problem of the discrete logarithm is reduced to the computation of the ME in the form (2). The discrete logarithm is considered to be a unidirectional function (1), because it is difficult to calculate it in a relatively acceptable time, for example, to break the cryptographic code. The development of an efficient computational algorithm for integer power of a modulo number for large numbers is relevant for solving problems of modern asymmetric cryptography, for the effective implementation of theoretical and numerical transformations and other applied problems. Therefore, it is very important to build algorithmic schemes that provide fast calculation of the ME.

2 REVIEW OF THE LITERATURE

Many effective methods of ME have been proposed [2, 3]. Among them are called: right-to-left k-ary exponentiation, left-to-right k-ary exponentiation, sliding window exponentiation, Montgomery ladder, simultaneous multiple exponentiation and their modifications. Considerable attention is paid to their software or hardware implementation [4–6] aimed at the effective definition of the discrete logarithm x.

One of the ways to accelerate the computation of modular elevation to the power is to parallelize calculations using modern technologies in universal computer systems [4–6].

Mathematical software libraries are used to implement the computation of ME. For example, the Pari/GP software library [7] contains a large set of programs for efficient computations of mathematical functions. The Pari/GP library also includes computation of the ME function for long numbers and other special numbers. A highly optimized modification of the well-known GMP or GNU Multiple Precision Arithmetic Library the MPIR library [8] contains the function of the realization the computation of ME. The library of cryptographic algorithms and schemes Crypto ++ is implemented in C ++ and fully supports 32 and 64-bit architectures of many operating systems and platforms [9]. The library contains a set of available primitives for theoretical and numerical operations, such as generation and verification of prime numbers, arithmetic over a finite field, operations on polynomials.

3 MATERIALS AND METHODS

The general-purpose exponentiation algorithms referred to as repeated square-and-multiply algorithms. The papers of Knuth [10], Bach and Shallit [11] describe the right-to-left binary exponentiation method. Cohen [12] provides a more comprehensive treatment of the right- to-left and left-to-right binary methods along with their generalizations to the k-ary method.

The central idea to calculate $A^x \bmod N$ is to use the binary representation of the exponent x

$$x = (x_{(k-1)} x_{(k-2)} \dots x_2 x_1 x_0)_b,$$

$$x = \sum_{i=0}^{k-1} 2^i x_i \quad \text{and } x_i \in \{0,1\}. \quad (3)$$

We write the exponent x as a set of m parts that are equal in binary length r. That is, the binary representation of the value of x consists of m, the bit length of each of them is equal to $r=k/m$. Then the binary representation of the exponent $\boldsymbol{x}$ will be

$$\begin{aligned} x &= (e_{(m-1)} \dots e_2 e_1)_b = \\ &(x_{mr-1} \dots x_{(m-1)(r+2)} x_{(m-1)(r+1)} x_{(m-1)r}) \dots \\ &(x_{2r-1} \dots x_{r+2} x_{r+1} x_r)(x_{r-1} \dots x_2 x_1 x_0)_b \end{aligned} \quad (4)$$

In this case, the x value will be

$$x = \sum_{i=0}^{k-1} 2^{i(k/m)} e_i. \quad (5)$$

Accordingly (4, 5), the computation of the ME takes the form

$$\begin{aligned} y = A^x \bmod N &= A^{(2^{(m-1)r} e_{(m-1)} \cdot \dots \cdot 2^{2r} e_2 \cdot 2^r e_1 \cdot 2^0 e_0)_b} \bmod N = \\ &= (A^{2^{(m-1)r} e_{(m-1)}} \bmod N * A^{2^{(m-2)r} e_{(m-2)}} \bmod N * \dots \\ &* A^{2^{2r} e_2} \bmod N * A^{2^r e_1} \bmod N * A^{2^0 e_0} \bmod N) \bmod N = \\ &= ((A^{e_{(m-1)}} \bmod N)^{2^{(m-1)r}} * (A^{e_{(m-2)}} \bmod N)^{2^{(m-2)r}} * \dots \\ &* (A^{e_2} \bmod N)^{2^{2r}} * (A^{e_1} \bmod N)^{2^r} * (A^{e_0} \bmod N)^{2^0}) \bmod N. \end{aligned} \quad (6)$$

There are three types of exponentiation algorithms $A^x \bmod N$ [13], which include:

1) basic techniques for exponentiation;
2) fixed-exponent x exponentiation algorithms;
3) fixed-base A exponentiation algorithms.

A fixed element of a group (generally z/qz) is repeatedly raised to many different powers in several cryptographic systems. A popular application of fixed-base exponentation is in elliptic curve cryptography, for instance for Diffie-Hellman key agreement and elliptic curve digital signature algorithm verification. Therefore, many research works have been focused on a fixed base of ME [14–16].

Compute, respectively (6), the value modulo N for a simple fixed-base A with exponents $x = 2^i = 1,2,4,8,16\dots$,

DOI 10.15588/1607-3274-2022-1-7

(i =0,1,2,…, r–1). Let A and N be relatively prime positive integers (A, N) = 1 and denote the least positive integer $x = \exp_N A$, in case

$$A^x \bmod N \equiv 1 \ . \qquad (7)$$

Accordance of the theorem [17], if A and N relatively prime (A, N) = 1, positive integer $\boldsymbol{x}$ is solution of the congruence (7) if and only if

$$x = q \cdot \exp_N A, \qquad (8)$$

Accordance the Euler's theorem, if A and N relatively prime (A, N) = 1, that $A^{\varphi(N)} \equiv 1 \pmod N$. Consequently, we can do conclusion

$$\varphi(N) = q \cdot \exp_N A, \qquad (9)$$

In case q=1, then $\varphi(N) = \exp_N R$, where R is the positive integer is called a primitive root modulo N. However the positive integer of modulo N, possesses a primitive root R if only if N=2, 4, P^k or $2P^k$, k is positive integer. The primitive root for modulo $N = P_1^{k1} P_2^{k2} \ldots P_m^{km}$ does not have, except крім if $\varphi(P_1^{k1})$, $\varphi(P_2^{k2})$,…, $\varphi(P_m^{ml})$ are relatively prime.

Thus, calculating $(R)^i \bmod N$ (i = 0,1, 2 ,,… N–1), we form a sequence of residuals ($r_0, r_1, r_2, \ldots, r_{i,\ldots}, r_{N-1}$), which periodically repeated for $x > (N-1)$ exponents. For all values of $A \in Zp$, the sequence $A^i \bmod P$ is cyclic for a non-primitive element.

The unique integer x with $1 \le x \le \varphi(N)$ and $R^x \bmod N \equiv A$ is called $\mathrm{ind}_R A$ index (or discrete logarithm) of A to base R modulo N. The properties of indeces, where a, b, k a positive integer and (a, N)=1, (b, N)=1, are

1) $\mathrm{ind}_R\ 1 \bmod \varphi(N) \equiv 0$,

2) $\mathrm{ind}_R (ab) \bmod \varphi(N) \equiv \mathrm{ind}_R (a) + \mathrm{ind}_R (b) \bmod \varphi(N)$,

3) $\mathrm{ind}_R (a^k) \bmod \varphi(N) \equiv k\ \mathrm{ind}_R (a) \bmod \varphi(N)$.

For example, for a primitive element R = 7, the sequence of residual values $r_i = (7^i) \bmod 11$,

$$(r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9) = (1,7,5,2,3,10,4,6,9,8).$$

The maximum period of repetitions is equal to $\exp_{11} 7 = 10$, because $7^{10} \bmod 11 = 1$, i=0,1,2,…,9. Then the sequence of indeces is equal (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) = = ($\mathrm{ind}_7 1$, $\mathrm{ind}_7 7$, $\mathrm{ind}_7 5$, $\mathrm{ind}_7 2$, $\mathrm{ind}_7 3$, $\mathrm{ind}_7 10$, $\mathrm{ind}_7 4$, $\mathrm{ind}_7 6$, $\mathrm{ind}_7 9$, $\mathrm{ind}_7 8$).

Accordingly of the property of indeces

1) $\mathrm{ind}_7\ 1 = 0$,

2) $\mathrm{ind}_7\ 6 = \mathrm{ind}_7\ (2*3) = \mathrm{ind}_7\ (2) + \mathrm{ind}_7\ 3 \bmod 10 = 3+4=7$;

3) $\mathrm{ind}_7\ 9 = \mathrm{ind}_7\ (3^2) = 2 * \mathrm{ind}_7\ 3 \bmod 10 = 2*4=8$.

In the case of calculating $(7^x) \bmod 11$ with index x = 32, the index will be equal to (32 mod ind117) = 2, and accordingly $\mathrm{ind}_7 5$. In the case of determining $(7^{2\,\wedge\,6}) \bmod 11$, we find the number of the residue in the sequence with the index $\mathrm{ind}_7 3$, which is equal to $(2^6) \bmod 10 = 4$. After all, the value of the ME for 2 elements in the sequence of residual values $r_4 = 3 = (7^{2\,\wedge\,6}) \bmod 11$.

For computations according to formula (6), we determine the residuals for exponents 2^i, (i =2,3,4,…). As a result of computations $r_i = (7^{2\,\wedge\,i}) \bmod 11$, ($i$ =2,3,4,…) we obtain the values of the residuals given in Table 1.

Table 1 – Periodic repetition of residual values $7 \wedge 2^i \bmod 11$

$7 \wedge 2^i$	7^0	7^1	7^2	7^4	7^8	7^{16}	7^{32}
T'=4 (**1**,**7**,5,3,9, 4)	1	7	**5**	**3**	**9**	**4**	5

7^{64}	7^{128}	7^{256}	7^{512}	7^{1024}	7^{2048}	7^{4096}
3	9	4	5	3	9	4

That is, in the process of computing $(7\wedge 2^i) \bmod 11$, starting with the exponent $2^1 = 2$, we obtain periodic repetition of the values of the residuals $r_0, r_1, \boldsymbol{r_2}, \boldsymbol{r_4}, \boldsymbol{r_8}, \boldsymbol{r_6}, r_2, r_4, r_8, r_{6,\ldots}$ with period T'= 4 and offset u = 0, because $2^0 = 1$.

The value of T' is found by the condition

$$A\wedge 2^i \bmod N \equiv A\wedge 2^{(i+T'+u)} \bmod N, i > u \ . \qquad (10)$$

Therefore, for a fixed-basis A of the ME of the computation of formula (6), which is equal to the product of the residuals of the exponent $(A\wedge 2^i) \bmod N$, (i = 2,3,4,…), you can speed up the process of computing the ME by precomputing the sequence of residuals what repetitions with the period T' after the offset u.

4 EXPERIMENTS

Mathematical software libraries are used to implement the computation of the ME. For example, the Pari/GP software library [7] contains a large set of programs for fast computations of mathematical functions. The Pari/GP library also includes computations of the Mod (a, n) ^ m function for multi-bit numbers, while using a small amount of memory in the process of performing computations. To work with numbers for modulo, the library uses a separate type *t_INTMOD*. Its feature is to represent the number in a special form (Montgomery reduction), which simplifies the computation of division by modulo. The Pari / GP library can be used in Linux or Mingw operating systems.

The library of cryptographic algorithms and schemes Crypto ++ is implemented in C ++ and fully supports 32 and 64-bit architectures of many operating systems and platforms [9]. The library contains a set of available primitives for theoretical and numerical operations, such as generation and verification of prime numbers, arithmetic over a finite field, operations on polynomials. Each of the Crypto ++ library primitives includes a function set.

The function mod_arithmetic.Exponentiate (*base_crypto*, *exp_crypto*) raising the number to the power by modulo. The result of the function is written to the variable *actual result crypto*, and the computation time is fixed and averaged with the output value "*crypto++ average time*" in nanoseconds.

Compared to the Pari/GP library, the well-known MPIR library [8] is easier in use and can be compiled in

DOI 10.15588/1607-3274-2022-1-7

Windows easily. Therefore, to implement the algorithm for computing the integer power of a number modulo, we used the MPIR library, which is written in C and assembler, and provides the ability to compile its functions in Visual Studio C ++. Accordingly, in the MPIR library, the data type *mpz t* represents large numbers of arbitrary length, which are selected for the power *exp* of the number *base* and the *mod* module with the number of bits from 256 to 2048 bits for testing.

The function mpz_powm (*expected_result*, *base*, *exp*, *mod*) performs raising the number to the power by modulo from the MPIR library, implementing the algorithm of the sliding window ("Sliding Window") with the use of Montgomery multiplication [14]. The result of the function is written to the variable *expected result*, and the computation time is fixed and averaged with the output value "*mpz powm average time"* in nanoseconds.

The function period_mod_exp (*remainders data, exp*) has been developed, which performs the basic iterative algorithm "Right-to-left binary exponentiation" [13]. To implement the algorithm, the library functions mpz_init_set (*mul, base*), mpz_sizeinbase (exp, 2), mpz_tstbit (*exp, i*), mpz_mul (*r, r, mul*) from the MPIR library are used, the parameters of which are multi-bit data up to 2048 bits. The algorithm is executed without dividing the exponent into parts, according to formulas (3–6) with m = 1, in one main stream. The function period_mod_exp () computes products modulo using precomputed residuals. The organization of the computation of the ME is performed respectively (11) and the scheme for computing A^x mod N in Fig. 1.

$$y = A^{x_{(k-1)}x_{(k-2)}\cdots x_2 x_1 x_0} \bmod N = \\ = (A^{2^{k-1}} \bmod N * A^{2^{k-2}} \bmod N * \ldots \\ * A^{2^2} \bmod N * A^{2^1} \bmod N * A^{2^0} \bmod N) \bmod N; \quad (11)$$

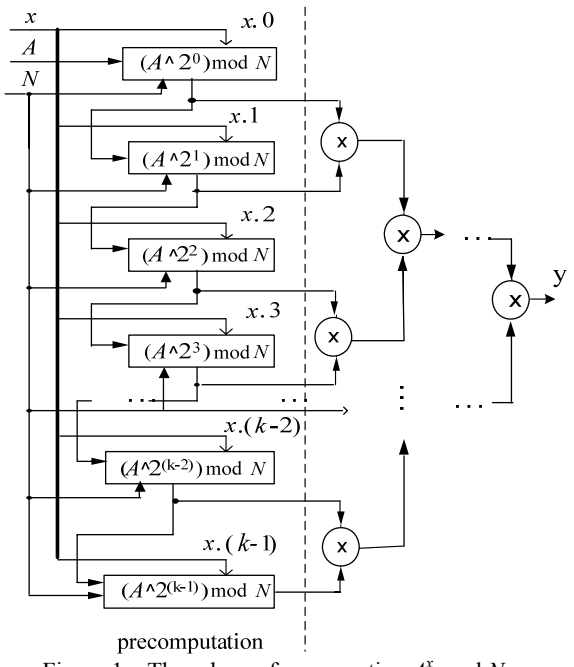


Figure 1 – The scheme for computing A^x mod N

In the software implementation, the function period_mod_exp (*remainders data, exp*) computes the products modulo (11) over the precomputed values of the residuals (A ^ 2^i) mod N, which are read using the function get_remainder (*const RemaindersData & data, size t power*). In the cycle of the function mpz_tstbit (exp, *i*) binary bits *x.i* of exponent *exp* are analyzed to determine to perform or not a multiplication operation modulo (Fig. 2). The computation of the value of the ME ends by writing the result in the variable *period_mod_exp_result*.

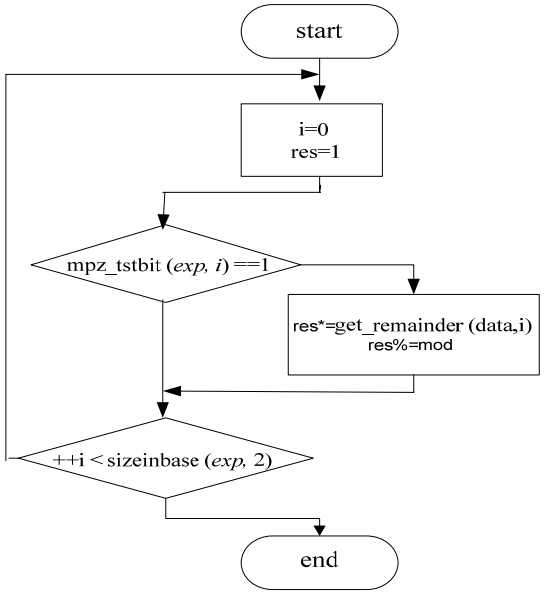


Figure 2 – The chart of the algorithm for determining to perform or not a multiplication under modulo in the function period_mod_exp() to compute the value of the ME

The precomputation includes finding the sequence of residuals for fixed numbers *Base* and *mod* for exp = 2^i (i = 0,1,2,…) and analysis of periodicity. In the program for computing the sequence of residuals is performed by the function find_remainders (*const mpz_class & base, const mpz_class & mod, size_t max_exp_bits*), which contains the function bool find_period (*const std::vector<mpz_class> & remainders*) to set the indication of finding the period. The function update_remainders (*RemaindersData & data*), shortens the length of the sequence of residuals to the end of the first periodicity. This function writes the offset *period_offset* beginning of the period and the length of the period *period_size* in the corresponding fields of the structure RemaindersData {*mpz_class base; mpz_class mod; std::vector <mpz_class> remainders; size_t period offset; size t period size;*} also.

The precomputation have been made in a separate function find_remainders () to optimize multiple residual searches (A^2^i) mod N. The peculiarity of the large values of *Base*, *mod* and *Exp* is also taken into account for which the residuals must be calculated, in case when the value of the period T' is many orders of magnitude greater than the number of bits of the Exp exponent.

DOI 10.15588/1607-3274-2022-1-7

5 RESULTS

To compare the computation efficiency of the developed ME function for a fixed basis with precomputation, two ME functions implemented from the Crypto ++ 8.2 and MPIR libraries are used. The comparison is performed with previously developed functions Single (), which performs in one main thread without taking into account the periodicity, and Parallel (), which performs in using two threads [18] computation of the ME.

Numerical experiments were carried out on a computer system with a multi-core microprocessor with shared memory in a 64-bit Windows. Testing was performed on computer systems with processors an Intel Core i9-10980XE (18 cores, 36 threads, 3.0GHz) and AMD Ryzen 3600(6 cores, 12 threads, 3.0GHz).

The average time data of the test with prime numbers P for *Base* and *mod*, that are

Base = P=131071,

Base = $P_1*P_2*P_3$=131080=8*5*3277,

mod =P =6700417, mod= P^2=$(5\ 039)^2$=25391521,

mod =P^3=$(5039)^3$=127947874319,

mod =$P_1*P_2*P_3$=(641*809*5039)=2613069191 are shown in Table 2.

Table 2 – The average execution time (ns) of the function period_mod() of computing the ME

Release/x86	Release/x86 Intel Core i9-10980XE, trials=2000					
Base *Exp* *mod*	131071 11039 263374721	131080 11039 263374721	131071 263375000 263374721	131080 263375000 263374721	131071 6039 127947874319	131080 6039 127947874319
period_mod()	745	762	1091	1162	710	798
Base *Exp* *mod*	131071 5039 6700417	131071 6039 6700417	131071 6700500 6700417	131071 11039 25391521	131080 11039 25391521	131071 26391521 25391521
period_mod()	791	833	1019	838	853	1110
Base *Exp* *mod*	131071 6700500 127947874319	131080 6700500 127947874319	131071 6039 2613069191	131080 6039 2613069191	131071 6700500 2613069191	131080 6700500 2613069191
period_mod()	1163	1133	729	731	1050	1060

To compute the ME with a given number of trials the values of exponent *Exp*, numbers *Base* and *mod* were given by pseudo-random numbers with number of binary digit to 2048 bits. To reduce the total computation time on increasing the number of digits of big numbers the number of trials of latch-up of the computation time is decrease correspondent. The results are presented in Table 3, which contains the values of average execution time (ns nanoseconds) of computing the ME for pseudo-random data *Base*1, *Exp*1, *mod1* for 1024 bits and *Base*2, *Exp*2, *mod*2 for 2048 bits, that are

Base1 =
15592587752839448261461062599367458801910077106635921807855458716257088123956757680446112611588790379930841
98450985791808156700960355218748709089617996382691960685601037523086698181288606777194603813043975878625936
07968358286567068579479763671817955144283945749615768573725580291910494735428411976050787788916;
Exp1 =
14283520978648999717087325550794657355644821408462852460542118822409968732298467354361417685207105354741569
82526225738456830754529824749225178413672786090891369885834477564794839184417957332157713350938927468516042
52279730368411597397577626119447392355080344631875081748708394736819710836176156337925349995164;
mod1 =
58915722462978682534126247597454067955853336194376986361024501907189851818542729349015243532285142254309917
03852776048028896537550292368120372032210211051014369063473209609243694640800275752961327153630792783722354
53227772400189542527413204742837529834451029227736537618930932834665884880224787395261044582 88;
and
Base2 =
25677387604979174745650113439241870319948692169987586987064217095280862955992909072300653168631621794286691
44090248166533311695144588834441618096640734511107106531362356071374321507249531854461586787197202959282597
81123638183830596292580376934671270834776657789712993784966788640286174086177056566697844654876749702991335
12869237149575978169492117082727320204008519907241837229067993684410038784610185215488903193461143855868161
08217151247348288474481211605784542061549242679745890886509283127487243351737251588531055149430134861136434
04436304687646801817005256929894904468321901418919444734072243769450781284602123 40;
Exp2 =
20697118667460294289329131926842862371260858718961622542390394128577965883462065464726476265621500584422101
09032488059047889420506452685343718712576517249128576248539133195446658184539707742058059362765132351678047
57330671171360229336910074202766840819523964084411554460264006668359867024199348668244418903647742281408991
58959010059757494492005981153055872187442495482411745134646120584804555929554183515697922429188623722060569
89487126897246500808974441045354232827662089593877770429717698803277620331558579804823032389188893339269852
03629914823135222855353547016859173882001780159272297308256145856605458847223736 80;

DOI 10.15588/1607-3274-2022-1-7

mod2 =
157517231345720355959956874368125960825444057365686170138216251141816860352630645141956911588687805530874096758773936173192212849470234982370978532269351866027326714684744912477506770434048787013558267810204995109644659341905468189951833441079292130714349995426535856054589395269550223482468472937086653958526469094233483743655909519384327711310830332518627465016808285004890531863472993853741749068729972978888527926301320033907702162996090456861888551577291792328064465975445931146310318328877160666812178649204722281454277435096606367577176097739534345883619710119588858725190093318844737746640231808576238875810 68.

Testing for the average execution time of computation of ME (Table 3) was performed by the functions: mpz_powm () from the MPIR library, crypto++ () from the Crypto ++ library. The comparison is performed with previously [22] developed functions Single () and Parallel (). The developed function period_mod () performs the computation of ME by forming an reduced sequence of residuals. The precomputation time to determine of the sequence of residuals is not taken into account.

Table 3 – The average execution time (ns)of the functions of computing the ME

Release/x86	AMD Ryzen 3600		Intel Core i9-10980XE	
Data	*Base1, Exp1, mod1*	*Base2,Exp2,mod2*	*Base1,Exp1, mod1*	*Base2, Exp2, mod2*
bits / trials	1024 / 1000	2048 / 500	1024 / 1000	2048 / 500
Single()	1993761	12916466	2032243	13445459
Parallel()	1678701	9129259	1938135	11366590
crypto++()	2484181	10668126	2607767	10915908
mpz_powm()	1167370	8264648	1196241	8969671
period_mod()	739048	4827014	724754	4927932

The results of the calculation of ME with all functions are compared for the accuracy of their implementation, which confirms the possibility of using the property of periodicity of the sequence of residuals for powers equal to integers of degree two.

6 DISCUSSION

The period_mod () function of the ME reduces the computation time relative to other functions with increasing bit size, starting from data values from 512 bits. Reducing the computation time of the period_mod () function as well as Single () and Parallel () depends on the number of logical one in the binary representation of the *Exp* exponent, which determines the number of multiplication operations in the main stream. The periodicity of the sequence of residues has its own characteristics and depends on the specific values of *Base*, *mod* and *Exp*, because they can differ by many orders of magnitude bits. In the Table 2 shows the cases when *Base* and *mod* are relatively prime (*Base*, *mod*) = 1. The results of the average execution time for the given relatively prime data are consistent with the basic properties that are well studied in number theory.

The software implementation period_mod () through a single-threaded computation shows a slight reduction in the time of determination of the modular exponent with an increase throughput of microprocessors (Table 3). Therefore, based on of the developed software the further implementation of the computation of ME using multithreaded technologies will provide an opportunity the efficient computation of discrete logarithm.

CONCLUSIONS

The work compares and analyses the developed software implementation of the computation of ME and the software implementation of the functions of Crypto++ and MPIR libraries. The computational scheme of the ME, the software implementation of the algorithm using single thread for computing of ME, the run time results of the computation on multi-core microprocessors of universal computer systems have been described. As a result, has developed the function period_mod() of the computation, what speedups the execution of the computations of ME for fixed-base with precomputation. The execution time of the algorithms depends on the specific values of the *Base, mod* and *Exp* of modular exponentiation. The software implementation with increasing the number of binary digits of data shows a reduction of computation time near two times with regard to the MPIR function of computing modular exponentiation.

The scientific novelty of obtained results lies in the implementation of the algorithm of computing the modular exponentiation based on the use of a reduced set of residuals and the fundamental property of modularity.

The practical significance of the work lies in the fact that the obtained results can be successfully apply in the modern asymmetric cryptography, for efficient computation of number-theoretic transforms and other computational problems.

Prospects for further research are that the developed function period_mod() can be used for the organization of multithreading computations of ME.

ACKNOWLEDGEMENTS

The authors are grateful to Roman Rykmas of the team leader of Uniservice LtdC for participation in testing and discussing the results obtained by the functions of computing the ME.

DOI 10.15588/1607-3274-2022-1-7

REFERENCES

1. Studholme C. The Discrete Log Problem [Electronic resource]. Department of Computer Science, University of Toronto, 2002, 57 p. Access mode: http://www.cs.toronto.edu/~cvs/dlog/research_paper.pdf
2. Jakubski A., Perliński R. Review of General Exponentiation Algorithms, *Scientific Research of the Institute of Mathematics and Computer Science*, 2011, Vol. 2, No. 10, pp. 87–98
3. Marouf I., Asad M. M., Al-Haija Q. A. Comparative Study of Efficient Modular Exponentiation Algorithms, *COMPUSOFT, An international journal of advanced computer technology,* August-2017, Vol. 6, Issue 8, pp. 2381–2392.
4. Lara P., Borges F., Portugal R., Nedjah N. Parallel modular exponentiation using load balancing without precomputation, *Journal of Computer and System Sciences*, 2012, Vol. 78, No. 2, pp. 575–582. https://doi.org/10.1016/j.jcss.2011.07.002
5. Nedjah N., Mourelle Ld. M. Three hardware architectures for the binary modular exponentiation: Sequential, parallel, and systolic, *Circuits and Systems I: Regular Papers, IEEE Transactions,* 2006. Vol. 53, Issue 3, pp. 627–633. https://doi.org/ 10.1109/TCSI.2005.858767.
6. Vollala S., Ramasubramanian N., Tiwari U. Energy-Efficient Modular Exponential Techniques for Public-Key Cryptography. Springer Nature, Singapur, Pte Ltd. 2021, 255 p. https://doi.org/10.1007/978-3-030-74524-0
7. PARI/GP home. [Electronic resource]. Access mode: http://pari.math.u-bordeaux.fr/
8. MPIR: Multiple Precision Integers and Rationals. [Electronic resource]. Access mode: http://mpir.org/
9. Crypto++ Library 8.6 Electronic resource]. Access mode: https://www.cryptopp.com
10. Knuth D. E. The art of computer programming. 3d ed. Reading (Mass), Addison-Wesley, cop. 1998, 712 p.
11. Bach E., Shallit J. Algorithmic Number Theory, Volume I, Efficient Algorithms. Cambridge, USA: MIT Press. 1996, 516 p.
12. Cohen H. A course in computational algebraic number theory. Berlin, Heidelberg, Springer. 1993, 536 p. https://doi.org/10.1007/978-3-662-02945-9
13. Menezes A. J., Oorschot van P. C., Vanstone S. A.. Handbook of Applied Cryptography, 5th printing, Boca Raton. CRC Press, 2001, 816 p.
14. Sorenson J. P. [Electronic resource] A sublinear-time parallel algorithm for integer modular exponentiation, 1999. pp. 1–8. Access mode: https://www.researchgate.net/publication/2274099_A
15. Robert J.-M., Negre C., Plantard T. Efficient Fixed Base Exponentiation and Scalar Multiplication based on a Multiplicative Splitting Exponent Recoding, *Journal of Cryptographic Engineering, Springer*, 2019, Vol. 9, Issue 2, pp. 115–136. https://doi.org/10.1007/s13389-018-0196-7.
16. Joye M. and Tunstall M. Exponent Recoding and Regular Exponentiation Algorithms, *Conference on Cryptology in Africa (Africacrypt 2009): Second International Conference*. Gammarth, Tunisia, 2009, proceedings. Published by Springer, 2009, pp. 334–349.
17. Rosen K. H. Elementary number theory and its applications 6th ed., China: Pearson/Addison Wesley, 2011, 721 p.
18. Prots'ko I. Kryvinska N., Gryshchuk O. The Runtime Analysis of Computation of Modular Exponentiation, *Radio Electronics, Computer Science, Control*, 2021, No. 3, pp. 42–47. DOI: https://doi.org/10.15588/1607-3274-2021-3-4

Received 20.12.2021.
Accepted 15.01.2022.

УДК 004.421

ОБЧИСЛЕННЯ МОДУЛЬНОЇ ЕКСПОНЕНТИ ДЛЯ ФІКСОВАНОЇ ОСНОВИ З ПЕРЕДОБЧИСЛЕННЯМ СКОРОЧЕНОГО НАБОРУ ЗАЛИШКІВ

Процько І. – д-р техн. наук, доцент, кафедра автоматизованих систем управління, Національний університет «Львівська політехніка», Львів, Україна.

Грищук О. – розробник програмного забезпечення, ТОВ «СофтСерв», Львів, Україна.

АНОТАЦІЯ

Актуальність. Модульне піднесення до степеня є важливою операцією в багатьох застосуваннях, що вимагає великої кількості обчислень. Швидкі обчислення модульної експоненти вкрай необхідні для ефективних обчислень у теоретично-числових перетвореннях, для забезпечення високої криптостійкості інформаційних даних та в багатьох інших завданнях.

Мета – аналіз часу виконання програмних функцій розрахунку модульної експоненти з розробленою програмою, що використовує попереднє обчислення зменшеного набору залишків для фіксованої бази.

Метод. Модульне піднесення до степеня реалізовано з використанням методу двійкового зсуву справа наліво для фіксованого базису з попереднім обчисленням зменшеного набору залишків. Для ефективного обчислення модульної експоненти великих чисел використовується властивість періодичності послідовності залишків фіксованої бази з експонентами, що дорівнюють цілочисельній степені двійки.

Результати. Проведено порівняння часу виконання п'яти варіантів функцій для обчислення модульного піднесення до степеня. В алгоритмі з попереднім обчисленням зменшеного набору залишків для фіксованої бази забезпечується більш швидке обчислення модульної експоненти для значень даних, що перевищують 1K двійкових розрядів, порівняно з функціями модульного піднесення до степеня бібліотек MPIR і Crypto++. Бібліотека MPIR з цілочисельним типом даних з кількістю двійкових розрядів від 256 до 2048 біт використовується для розробки алгоритму обчислення модульного піднесення до степеня.

Висновки. У роботі розглянуто та проаналізовано розроблену програмну реалізація обчислення модульної експоненти на універсальних комп'ютерних системах. Одним із способів реалізації прискорення обчислення модульного піднесення до степеня є розробка алгоритмів, які можуть використовувати попереднє обчислення зменшеного набору залишків для фіксованої бази. Програмна реалізація модульного піднесення до степеня зі збільшенням від числа 1K двійкових розрядів даних показує покращення часу обчислень у порівнянні з функцією модульного піднесення до степеня бібліотек MPIR та Crypto++.

КЛЮЧОВІ СЛОВА: модульне піднесення до степеня, великі числа, алгоритм зведення до степеня, фіксована базова ступінь, множина залишків.

DOI 10.15588/1607-3274-2022-1-7

УДК 004.421

ВЫЧИСЛЕНИЕ МОДУЛЬНОЙ ЭКСПОНЕНТЫ ДЛЯ ФИКСИРОВАННОЙ ОСНОВЫ С ПРЕДВЫЧИСЛЕНИЕМ СОКРАЩЕННОГО НАБОРА ОСТАТКОВ

Процько И. – д-р техн. наук, доцент кафедры автоматизированных систем управления, Национальный университет «Львивська политэхника», Львов, Украина.

Грищук О. – разработчик программного обеспечения, ООО «СофтСерв», Львов, Украина.

АННОТАЦИЯ

Актуальность. Возведение в степень – важная операция во многих приложениях, требующая большого количества вычислений. Быстрые вычисления модульного возведения в степень необходимы для эффективных вычислений в теоретико-численных преобразованиях, для обеспечения высокой криптостойкости информационных данных и во многих других приложениях.

Цель – анализ времени выполнения программных функций расчета модульной экспоненты с разработанной программой, использующей предварительные вычисления сокращенного набора остатков для фиксированной базы.

Метод. Модульное возведение в степень реализовано с использованием разработки метода двоичного сдвига справа налево для фиксированного базиса с предварительным вычислением уменьшенного набора остатков. Для эффективного вычисления модульной экспоненты больших чисел используется свойство периодичности последовательности остатков фиксированной базы с экспонентами, равными целочисленной степени двойки.

Результаты. Проведено сравнение времени выполнения пяти вариантов функций для вычисления модульной экспоненты. В алгоритме с предварительным вычислением сокращенного остатка набор для фиксированной базы обеспечивается более быстрое вычисление модульного возведения в степень для значений, превышающих 1К двоичных цифр, по сравнению с функциями модульной экспоненты библиотек MPIR и Crypto++. Библиотека MPIR с целочисленным типом данных с количеством двоичных разрядов от 256 до 2048 бит используется для разработки алгоритма вычисления модульного возведения в степень.

Выводы. В работе рассмотрена и проанализирована разработанная программная реализация вычисления модульной экспоненты на универсальных компьютерных системах. Один из способов реализации ускорения вычисления модульного возведения в степень является разработка алгоритмов, которые могут использовать предварительное вычисление сокращенного набора остатков для фиксированной базы. Программная реализация модульного возведения в степень с увеличением с 1024 числа двоичных разрядов экспоненты показывает улучшение времени вычислений по сравнению с функциями модульной экспоненты библиотек MPIR и Crypto++.

КЛЮЧЕВЫЕ СЛОВА: модульное возведение в степень, большие числа, алгоритм возведения в степень, возведение в степень с фиксированной основой, множество остатков.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Studholme C. The Discrete Log Problem [Electronic resource] / C. Studholme // Department of Computer Science, University of Toronto. – 2002. – 57 p. Access mode: http://www.cs.toronto.edu/~cvs/dlog/research_paper.pdf
2. Jakubski A. Review of General Exponentiation Algorithms / A. Jakubski, R. Perliński // Scientific Research of the Institute of Mathematics and Computer Science. – 2011. – Vol. 2, № 10. – P. 87–98
3. Marouf I. Comparative Study of Efficient Modular Exponentiation Algorithms. / I. Marouf, M. M. Asad, Q. A. Al-Haija // COMPUSOFT, An international journal of advanced computer technology. – August-2017. – Vol. 6, Issue 8. – P. 2381–2392.
4. Parallel modular exponentiation using load balancing without precomputation / [P. Lara, F. Borges, R. Portugal, N. Nedjah] // Journal of Computer and System Sciences. – 2012. – Vol. 78, № 2. – P. 575–582. https://doi.org/10.1016/j.jcss.2011.07.002
5. Nedjah N. Three hardware architectures for the binary modular exponentiation: Sequential, parallel, and systolic / N. Nedjah, Ld. M. Mourelle // Circuits and Systems I: Regular Papers, IEEE Transactions. – 2006. – Vol. 53, Issue 3. – P. 627–633. https://doi.org/ 10.1109/TCSI.2005.858767.
6. Vollala S. Energy-Efficient Modular Exponential Techniques for Public-Key Cryptography. / S. Vollala, N. Ramasubramanian, U. Tiwari. – Springer Nature, Singapur: Pte Ltd. 2021. – 255 p. https://doi.org/10.1007/978-3-030-74524-0
7. PARI/GP home. [Electronic resource]. – Access mode: http://pari.math.u-bordeaux.fr/
8. MPIR: Multiple Precision Integers and Rationals. [Electronic resource]. – Access mode: http://mpir.org/
9. Crypto++ Library 8.6 Electronic resource]. – Access mode: https://www.cryptopp.com
10. Knuth D. E. The art of computer programming / D. E. Knuth. – 3d ed. Reading (Mass): Addison-Wesley, cop. 1998. – 712 p.
11. Bach E. Algorithmic Number Theory / E. Bach, J. Shallit. – Volume I: Efficient Algorithms. – Cambridge, USA: MIT Press. 1996. – 516 p.
12. Cohen H. A course in computational algebraic number theory / H. Cohen. – Berlin, Heidelberg : Springer, 1993. – 536 p. https://doi.org/10.1007/978-3-662-02945-9
13. Menezes A.J. Handbook of Applied Cryptography / A. J. Oorschot, S. A. Vanstone. – 5th printing, Boca Raton : CRC Press. 2001. – 816 p.
14. Sorenson J. P. [Electronic resource] A sublinear-time parallel algorithm for integer modular exponentiation, 1999. – P. 1–8. – Access mode: https://www.researchgate.net/publication/2274099
15. Robert J.-M. Efficient Fixed Base Exponentiation and Scalar Multiplication based on a Multiplicative Splitting Exponent Recoding / J.-M. Robert, C. Negre, T. Plantard // Journal of Cryptographic Engineering, Springer. – 2019. – Vol. 9, Issue 2. – P. 115–136. https://doi.org/10.1007/s13389-018-0196-7.
16. Joye M. Exponent Recoding and Regular Exponentiation Algorithms / M. Joye and M. Tunstall // Conference on Cryptology in Africa (Africacrypt 2009): Second International Conference, Gammarth, Tunisia, 2009: proceedings. – Published by Springer, 2009. –P. 334–349.
17. Rosen K. H. Elementary number theory and its applications / K. H. Rosen. – 6th ed., China : Pearson/Addison Wesley, 2011. – 721 p.
18. Prots'ko I. The Runtime Analysis of Computation of Modular Exponentiation / I. Prots'ko, N. Kryvinska, O. Gryshchuk / Радіоелектроніка, інформатика, управління. – 2021. – № 3. – С. 42–47. DOI: https://doi.org/10.15588/1607-3274-2021-3-4

DOI 10.15588/1607-3274-2022-1-7

UDC 621.391.1: 621.395

TWO PAIRS OF DUAL QUEUEING SYSTEMS WITH CONVENTIONAL AND SHIFTED DISTRIBUTION LAWS

Tarasov V. N. – Dr. Sc., Professor, Head of Department of Software and Management in Technical Systems of Volga State University of Telecommunications and Informatics, Samara, Russian Federation.

Bakhareva N. F. – Dr. Sc., Professor, Head of Department of Informatics and Computer Engineering of Volga State University of Telecommunications and Informatics, Samara, Russian Federation.

ABSTRACT

Context. The relevance of studies of G/G/1 systems is associated with the fact that they are in demand for modeling data transmission systems for various purposes, as well as with the fact that for them there is no final solution in the general case. We consider the problem of deriving a solution for the average delay of requests in a queue in a closed form for ordinary systems with Erlang and exponential input distributions and for the same systems with distributions shifted to the right.

Objective. Obtaining a solution for the main characteristic of the system – the average delay of requests in a queue for two pairs of queuing systems with ordinary and shifted Erlang and exponential input distributions, as well as comparing the results for systems with normalized Erlang distributions.

Methods. To solve the problem posed, the method of spectral solution of the Lindley integral equation was used, which allows one to obtain a solution for the average delay for the systems under consideration in a closed form. For the practical application of the results obtained, the method of moments of the theory of probability was used.

Results. Spectral solutions of the Lindley integral equation for two pairs of systems are obtained, with the help of which calculation formulas are derived for the average delay of requests in the queue in a closed form. Comparison of the results obtained with the data for systems with normalized Erlang distributions confirms their identity.

Conclusions. The introduction of the time shift parameter into the distribution laws of the input flow and service time for the systems under consideration transforms them into systems with a delay with a shorter waiting time. This is because the time shift operation reduces the value of the variation coefficients of the intervals between the arrivals of claims and their service time, and as is known from the queuing theory, the average delay of requests is related to these variation coefficients by a quadratic dependence. If a system with Erlang and exponential input distributions works only for one fixed pair of values of the coefficients of variation of the intervals between arrivals and their service time, then the same system with shifted distributions allows operating with interval values of the coefficients of variations, which expands the scope of these systems. The situation is similar with shifted exponential distributions. In addition, the shifted exponential distribution contains two parameters and allows one to approximate arbitrary distribution laws using the first two moments. This approach makes it possible to calculate the average latency and higher-order moments for the specified systems in mathematical packets for a wide range of changes in traffic parameters. The method of spectral solution of the Lindley integral equation for the systems under consideration has made it possible to obtain a solution in closed form, and these obtained solutions are published for the first time.

KEYWORDS: Erlang and exponential distribution laws, Lindley integral equation, spectral expansion solution method, Laplace transform.

ABBREVIATIONS

LIE is a Lindley integral equation;

QS is a queuing system;

PDF is a probability distribution function.

NOMENCLATURE

$a(t)$ is a density function of the distribution of time between arrivals;

$A^*(s)$ is a Laplace transform of the function $a(t)$;

$b(t)$ is a density function of the distribution of service time;

$B^*(s)$ is a Laplace transform of the function $b(t)$;

c_λ the coefficient of variation of time between arrivals;

c_μ the coefficient of variation of service time;

E_2 is a ordinary Erlang distribution of the second order;

E_2^- is a shifted Erlang distribution of the second order;

G is a arbitrary distribution law;

M is a exponential distribution law;

M^- is a shifted exponential distribution law;

$\overline{W}$ is a average waiting time in the queue;

$W^*(s)$ is a Laplace transform of waiting time density function;

λ is a Erlang (exponential) distribution parameter for input flow;

μ is a Erlang (exponential) distribution parameter for of service time;

ρ is a system load factor;

$\overline{\tau}_\lambda$ is a average time between arrivals;

$\overline{\tau}_\lambda^2$ is a second initial moment of time between arrivals;

$\overline{\tau}_\mu$ is a average service time;

$\overline{\tau_\mu^2}$ is a second initial moment of service time;

$\Phi_+(s)$ is a Laplace transform of the PDF of waiting time;

DOI 10.15588/1607-3274-2022-1-8

$\psi_+(s)$ is a first component of spectral decomposition;

$\psi_-(s)$ is a second component of spectral decomposition.

INTRODUCTION

This article is devoted to the analysis of two pairs of QSs, including the Erlang distribution law as a special case of a more general gamma distribution law and an exponential distribution. The task is to derive solutions for the average delay of requests in the queue, which is the main characteristic for any QS. This characteristic, for example, is used to estimate packet delays in packet-switched networks when they are modeled using QS. The considered QSs, according to the three-position symbolism introduced by Kendall for their classification, we denote by $E_2/M/1$ and $M/E_2/1$. Here the distribution law for E_2 differs from the previously considered normalized Erlang distribution.

We also investigated the above systems with time-shifted input distributions in order to obtain a solution for the average delay. In queuing theory, studies of G/G/1 systems are especially relevant because there is no solution in the final form for the general case and one has to carry out research for special cases of distribution laws. In the study of G/G/1 systems, an important role is played by the method of spectral solution of the Lindley integral equation [1]. The paper proposes new models of queuing with shifted second-order Erlang distributions, as a special case of the Gamma distribution law.

In the previous works of the authors [2–7], it was noted that the shift of the distribution laws in the QS by the value $t_0 > 0$ leads to a decrease in the average delay of requests in the queue due to a decrease in the coefficients of variation of the time intervals of arrivals c_λ and servicing c_μ. It is known that the average delay is related to these coefficients of variation by a quadratic dependence [1].

The object of study is the queueing systems type G/G/1.

The subject of study is the average queue delay in conventional systems $E_2/M/1$ and $M/E_2/1$ and in the same systems, but with shifted input distributions.

The purpose of the work is to obtain a solution in a closed form for the main characteristic of the system – the average delay in the queue for the above QS.

1 PROBLEM STATEMENT

The paper poses the problem of finding a solution for the delay of requests in the queue in conventional QS systems $E_2/M/1$ and $M/E_2/1$ and in the same QS with shifted input distributions. Here E_2 means the second-order Erlang distribution as a special case of a more general Gamma distribution law and has the form $f_\lambda(t) = \lambda^2 t e^{-\lambda t}$ for describing the distribution density of the arrival intervals, in contrast to the normalized distribution with the density function $f_\lambda(t) = 4\lambda^2 t e^{-2\lambda t}$. Moreover, these density functions differ in numerical characteristics.

In a brief presentation of the method of spectral expansion of the LIE solution, we will adhere to the approach and symbols of the author of the classics of the queuing theory [1]. At the heart of the LIE solution by the spectral expansion method is to find for expression $F_\lambda^*(-s)F_\mu^*(s)-1$ a representation in the form of a product of two factors, which would give a rational function of *s*. Consequently, to find the distribution law of the delay of requests in the queue, the following spectral expansion $F_\lambda^*(-s)F_\mu^*(s)-1 = \psi_+(s)/\psi_-(s)$ is necessary. Here $\psi_+(s)$ and $\psi_-(s)$ are some rational functions of *s* that can be factorized. Functions and must satisfy special conditions according to [1].

To solve this problem, it is first necessary to construct spectral solutions of the form $F_\lambda^*(-s)F_\mu^*(s)-1 = \psi_+(s)/\psi_-(s)$ for these systems, considering special conditions in each case.

2 REVIEW OF THE LITERATURE

The method of spectral decomposition of the solution of the Lindley integral equation used in this work is presented in detail for the first time in the classics of the queuing theory [1]. This method was used by the authors in [2–7] and in many other works in the study of QS with shifted distributions. The spectral solution method is widely used not only in queuing theory, but also in mathematics, physics, electromagnetism and other fields [8–12]. In both foreign and Russian-language literature, the authors have not found research results in this subject area.

The closest to this area are works [15, 16], where the questions of accessing Internet web resources as queues with time lag, described by Wiener-Hopf processes, are investigated.

The problems of approximating distribution laws using several initial moments of time intervals are covered in [11–14], and the results of new research in the queuing theory [18–27].

3 MATERIALS AND METHODS

As you know, the two-parameter gamma distribution is given by the density function of the form

$$f(t) = \begin{cases} \dfrac{\beta^{-\alpha} t^{\alpha-1} e^{-t/\beta}}{\Gamma(\alpha)}, & t \geq 0, \\ 0, & t < 0, \end{cases}$$

where $\Gamma(\alpha)$ is a gamma function equal $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ for any real number $z>0$, $\alpha>0$, $\beta>0$. In the case of integers, this distribution turns into an Erlang distribution of order α. For example, when replacing $\lambda = 1/\beta, k = \alpha$, we get the usual Erlang distribution of order *k*:

DOI 10.15588/1607-3274-2022-1-8

$$f_\lambda(t)=\frac{\lambda^k t^{k-1}e^{-\lambda t}}{(k-1)!}.$$

For a second-order distribution, the density function has the form $f_\lambda(t)=\lambda^2 te^{-\lambda t}$. This distribution differs from the previously considered normalized Erlang distribution, where $f_\lambda(t)=4\lambda^2 te^{-2\lambda t}$. The Erlang distribution was normalized in order to make the mathematical expectation independent of the order of the distribution *k*, therefore, the numerical characteristics of the two forms of writing the distribution will change.

The main differences between the normal (gamma-derived) and normalized Erlang distributions E_2 are shown below. Differences in numerical characteristics:

– for the usual Erlang distribution

$$\overline{\tau}_\lambda=2/\lambda,\ \overline{\tau_\lambda^2}=6/\lambda^2,\ c_\lambda^2=1/2,$$

– for normalized distribution

$$\overline{\tau}_\lambda=1/\lambda,\ \overline{\tau_\lambda^2}=3/(2\lambda^2),\ c_\lambda^2=1/2.$$

Differences in the distribution parameter obtained by the method of moments:

– for the usual Erlang distribution $\lambda=2/\overline{\tau}_\lambda$,

– for normalized distribution $\lambda=1/\overline{\tau}_\lambda$.

Thus, the indicated distribution laws differ in both parameter and numerical characteristics, except for the coefficient of variation. As we will see below, systems formed by ordinary and normalized Erlang distributions will have different spectral expansions. Due to such a difference between the distributions, we are interested in the fact whether this difference will affect the final result of the QS – the average delay of requests in the queue, especially in the case of shifted distributions.

In this regard, it will be interesting to see the results obtained.

Next, consider a system E_2/M/1 formed by two flows given by functions of distribution densities:

– for the input flow

$$f_\lambda(t)=\lambda^2 te^{-\lambda t}, \quad (1)$$

– for service times

$$f_\mu(t)=\mu e^{-\mu t}. \quad (2)$$

Let us write the Laplace transform of functions (1) and (2):

$$F_\lambda^*(s)=\left(\frac{\lambda}{\lambda+s}\right)^2,\ F_\mu^*(s)=\frac{\mu}{\mu+s}.$$

Then the spectral expansion of the LIE solution for the E_2/M/1 system takes the form:

$$\begin{aligned}F_\lambda^*(-s)F_\mu^*(s)-1&=\frac{\psi_+(s)}{\psi_-(s)}=\left(\frac{\lambda}{\lambda\text{-}s}\right)^2\frac{\mu}{\mu+s}-1=\\&=\frac{\lambda^2\mu-(\lambda-s)^2(\mu+s)}{(\lambda-s)^2(\mu+s)}=\\&=-\frac{s[s^2+(\mu-2\lambda)s+\lambda(\lambda-2\mu)]}{(\lambda-s)^2(\mu+s)}=\\&=-\frac{s(s+s_1)(s-s_2)}{(\lambda-s)^2(\mu+s)},\end{aligned} \quad (3)$$

since the quadratic equation $s^2+(\mu-2\lambda)s+\lambda(\lambda-2\mu)=0$ obtained from the expansion numerator has one negative root $-s_1=-(\mu-2\lambda)/2-\sqrt{\mu(\mu+4\lambda)}/2$ and one positive root $s_2=(2\lambda-\mu)/2+\sqrt{\mu(\mu+4\lambda)}/2$ in the case of a stable system with $\lambda<\mu$.

Therefore, we will take an expression $\psi_+(s)=\frac{s(s+s_1)}{s+\mu}$ as a function $\psi_+(s)$, since its zeros $s=0$, $s=-s_1$ and the pole $s=-\mu$ lie in the region $\mathrm{Re}(s)\le 0$, and we take an expression $\psi_-(s)=-\frac{(\lambda-s)^2}{s-s_2}$ as a function $\psi_-(s)$.

Finally, the components of the spectral expansion for the E_2/M/1 system will have the form

$$\psi_+(s)=\frac{s(s+s_1)}{s+\mu};\ \psi_-(s)=-\frac{(\lambda-s)^2}{s-s_2}. \quad (4)$$

Fig. 1 confirms the fulfillment of special conditions [1] where the zeros and poles of the fractional rational function on the complex s – plane are displayed to eliminate errors in constructing the spectral decomposition.

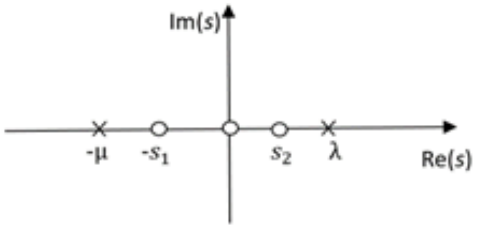


Figure 1 – Zeros and poles of the function $\psi_+(s)/\psi_-(s)$ for the E_2/M/1 system

In Fig. 1, the poles are marked with crosses, and the zeros are marked with circles.

Further, using the method of spectral decomposition, we find the constant *K*, which determines the probability that a demand entering the system finds it free:

$$K=\lim_{|s|\to 0}\frac{\psi_+(s)}{s}=\lim_{|s|\to 0}\frac{s+s_1}{s+\mu}=\frac{s_1}{\mu}.$$

DOI 10.15588/1607-3274-2022-1-8

Let us construct a function $\Phi_+(s)=\frac{K}{\psi_+(s)}=\frac{s_1(s+\mu)}{\mu s(s+s_1)}$ through which we find the Laplace transform of the delay density function:

$$W^*(s)=s\cdot\Phi_+(s)=\frac{s_1(s+\mu)}{\mu(s+s_1)}. \quad (5)$$

Derivative of a function $W^*(s)$ with a minus sign incl. $s=0$:

$$-\frac{dW^*(s)}{ds}|_{s=0}=-\left[\frac{s_1\mu(s+s_1)-s_1(s+\mu)\mu}{\mu^2(s+s_1)^2}\right]|_{s=0}=$$
$$=\frac{s_1\mu^2-s_1^2\mu}{\mu^2 s_1^2}=\frac{1}{s_1}-\frac{1}{\mu}.$$

Then the average delay of requests in the queue for the $E_2/M/1$ system:

$$\overline{W}=1/s_1-1/\mu, \quad (6)$$

where $s_1=(\mu-2\lambda)/2+\sqrt{\mu(\mu+4\lambda)}/2$ is the absolute value of the negative root $-s_1$. After the expression for the average waiting time for the $E_2/M/1$ system has been found, we can proceed to the study of the $E_2/M/1$ system with a time lag.

We denote such a system $E_2^-/M^-/1$. For this system, the distributions of the arrival and service intervals are described by the following shifted density functions:

$$f_\lambda(t)=\lambda^2(t-t_0)e^{-\lambda(t-t_0)}, \quad (7)$$

$$f_\mu(t)=\mu e^{-\mu(t-t_0)}. \quad (8)$$

Statement 1. The spectral expansion of the LIE solution for the $E_2^-/M^-/1$ system and the final formula for the average delay have exactly the same form as for the $E_2/M/1$ system, but with changed parameters due to a shift in the distribution laws.

Proof. Laplace transforms of functions (7) and (8) have the form:

$$F_\lambda^*(s)=\left(\frac{\lambda}{\lambda+s}\right)^2 e^{-t_0 s},\ F_\mu^*(s)=\frac{\mu}{\mu+s}e^{-t_0 s}.$$

For the $E_2^-/M^-/1$ system, the spectral expansion will have the form:

$$F_\lambda^*(-s)F_\mu^*(s)-1=\left(\frac{\lambda}{\lambda-s}\right)^2 e^{t_0 s}\times\frac{\mu}{\mu+s}e^{-t_0 s}-1=$$
$$=\left(\frac{\lambda}{\lambda-s}\right)^2\times\frac{\mu}{\mu+s}-1.$$

Here, the powers of the exponentials in the spectral decomposition are also zeroed, and thus the time shift operation is leveled. The last expression after the transformations will result in the result (3). Thus, the spectral expansions of the LIE solution for both systems coincide. Consequently, all the above calculations for the $E_2/M/1$ system are also valid for the $E_2^-/M^-/1$ system. Statement 1 is proved.

To determine the unknown distribution parameters E_2^-, we use the Laplace transform of function (7). The average value of the interval between arrivals is given by the first derivative of the Laplace transform with a minus sign at the point s=0:

$$-\frac{dF_\lambda^*(s)}{ds}|_{s=0}=\left[\frac{2\lambda^2 e^{-t_0 s}}{(\lambda+s)^3}+\frac{\lambda^2 t_0 e^{-t_0 s}}{(\lambda+s)^2}\right]|_{s=0}=2/\lambda+t_0.$$

From here

$$\overline{\tau}_\lambda=2/\lambda+t_0. \quad (9)$$

The second initial moment of the interval between arrivals is equal to

$$\left.\frac{d^2F_\lambda^*(s)}{ds^2}\right|_{s=0}=\frac{6}{\lambda^2}+4\frac{t_0}{\lambda}+t_0^2.$$

From here

$$\overline{\tau_\lambda^2}=\frac{6}{\lambda^2}+4\frac{t_0}{\lambda}+t_0^2.$$

Determine the square of the coefficient of variation

$$c_\lambda^2=\frac{\overline{\tau_\lambda^2}-(\overline{\tau_\lambda})^2}{(\overline{\tau_\lambda})^2}=\frac{2}{(2+\lambda t_0)^2}.$$

From here

$$c_\lambda=\sqrt{2}/(2+\lambda t_0). \quad (10)$$

Note that for the distribution E_2: $\overline{\tau}_\lambda=2/\lambda$, $c_\lambda=1/\sqrt{2}$. Consequently, because of the shift of the distribution laws by the value $t_0>0$, the coefficient of variation c_λ for the distribution E_2^- decreases by $(1+\lambda t_0/2)$ a factor of comparison with c_λ for the distribution E_2.

It remains to determine the numerical characteristics for the shifted exponential distribution M^-.

$$-\left.\frac{dF_\mu^*(s)}{ds}\right|_{s=0}=-\frac{d}{ds}\left[\left(\frac{\mu}{\mu+s}\right)e^{-t_0 s}\right]|_{s=0}=$$
$$=\left.\left[\frac{\mu t_0 e^{-t_0 s}(s+\mu)+\mu e^{-t_0 s}}{(s+\mu)^2}\right]\right|_{s=0}=1/\mu+t_0.$$

From here

DOI 10.15588/1607-3274-2022-1-8

$$\overline{\tau}_\mu = 1/\mu + t_0 \,. \qquad (11)$$

Using the second derivative of the Laplace transform for s=0, we define the second initial moment of service time

$$\left.\frac{d^2 F_\mu^*(s)}{ds^2}\right|_{s=0} = \frac{2}{\mu^2} + 2\frac{t_0}{\mu} + t_0^2 \,.$$

From here

$$c_\mu = 1/(1+\mu t_0) \,. \qquad (12)$$

Now setting the values obtained above $\overline{\tau}_\lambda$, $\overline{\tau}_\mu$, c_λ , c_μ as input parameters for calculating for the $E_2^-/M^-/1$ system, as well as the shift parameter t_0, you can calculate the average delay using formula (6). Here, the ranges of variation of the variation coefficients $c_\lambda \in \left(0, 1/\sqrt{2}\right)$ and $c_\mu \in (0,1)$, are determined by relations (10) and (12), respectively, depending on the magnitude of the shift parameter $0 < t_0 < \overline{\tau}_\mu$.

Next, consider the M/E_2/1 system formed by two flows given by the functions of the distribution densities of the intervals:

– for the input flow

$$f_\lambda(t) = \lambda e^{-\lambda t} \,, \qquad (13)$$

– for service time

$$f_\mu(t) = \mu^2 t e^{-\mu t} \,. \qquad (14)$$

The spectral solution of the Lindley integral equation for this system takes the form

$$F_\lambda^*(-s)F_\mu^*(s) - 1 = \frac{\lambda}{\lambda - s}\cdot\left(\frac{\mu}{\mu + s}\right)^2 - 1 = \frac{s\left[s^2 + (2\mu - \lambda)s + \mu(\mu - 2\lambda)\right]}{(\lambda - s)(\mu + s)^2} \,.$$

The square trinomial $s^2 + (2\mu - \lambda)s + \mu(\mu - 2\lambda)$ in the numerator of the expansion in the case of a stable system has two real negative roots $-s_1$, $-s_2$:

$$-s_1 = -(2\mu - \lambda)/2 + \sqrt{\lambda(\lambda + 4\mu)}/2 \,,$$

$$-s_2 = -(2\mu - \lambda)/2 - \sqrt{\lambda(\lambda + 4\mu)}/2 \,.$$

The final spectral solution will have the form

$$\frac{\psi_+(s)}{\psi_-(s)} = \frac{s(s+s_1)(s+s_2)}{(\lambda - s)(\mu + s)^2} \,.$$

Based on the rules for constructing functions $\psi_+(s)$ and $\psi_-(s)$ choose $\psi_+(s) = \frac{s(s+s_1)(s+s_2)}{(\mu+s)^2}$, $\psi_-(s) = \lambda - s$. The zeros and poles of this expansion are shown in Fig. 2, where the poles are marked with crosses and zeros are marked with circles.

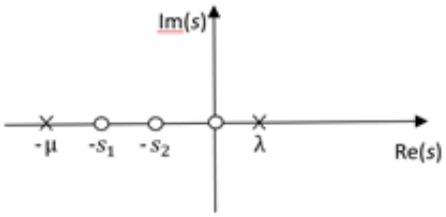


Figure 2 – Zeros and poles of the function $\psi_+(s)/\psi_-(s)$ for the M/E_2/1 system

The constant required to obtain a solution $K = \lim_{s\to 0}\frac{\psi_+(s)}{s} = \frac{s_1 s_2}{\mu^2}$. Next, we construct the function $\Phi_+(s) = \frac{K}{\psi_+(s)} = \frac{(1-\rho)(\mu+s)^2}{s(s+s_1)(s+s_2)}$. Whence it follows that the Laplace transform of the density function of the delay time in the M/E_2/1 system

$$W^*(s) = s\cdot\Phi_+(s) = \frac{(1-\rho)(\mu+s)^2}{(s+s_1)(s+s_2)} \,. \qquad (15)$$

The first derivative of function (14) with a minus sign at point s=0 is

$$-\frac{dW^*(s)}{ds}\Big|_{s=0} = \frac{1}{s_1} + \frac{1}{s_2} - \frac{2}{\mu} \,.$$

Hence the average delay of requests in the queue

$$\overline{W} = \frac{1}{s_1} + \frac{1}{s_2} - \frac{2}{\mu} \,. \qquad (16)$$

Comment. The Laplace transforms of the delay density functions (5) and (15) make it possible to obtain formulas not only for the average values of the delay, but also for the moments of higher orders for the delay. Considering the definition of jitter for telecommunications as the spread of the delay around the average value, then jitter can also be determined through the variance of the delay [17].

There is another way to obtain the formula for the average delay for the M/E_2/1 system. Because this system belongs to the class of M/G/1 systems, we will use the known result for this system by the Polyachek – Khinchin equation for the Laplace transform of the density function of the delay for the M/G/1 system [1]:

DOI 10.15588/1607-3274-2022-1-8

$$W^*(s)=\frac{s(1-\rho)}{s-\lambda+\lambda F_\mu^*(s)}, \quad (17)$$

where $F_\mu^*(s)=\mu^2/(s+\mu)^2$ is the Laplace transform of the service time density function.

The Polyachek-Khinchin formula [1], gives the average delay of requests in the queue in the M/G/1 system:

$$\bar{W}=\frac{\lambda\,\overline{\tau_\mu^2}}{2(1-\rho)}, \quad (18)$$

where $0<\rho=\lambda/\mu<1$. For the distribution E_2, the second initial moment of service time, then from (18) we obtain the average delay in the M/G/1 system:

$$\bar{W}=\frac{3\rho}{2\mu(1-\rho)}. \quad (19)$$

Now it remains to verify that equalities (16) and (19) are identical. When substituting already calculated values s_1, s_2 in (16), and performing simple mathematical calculations, we get a complete coincidence with formula (19).

Let us begin to determine the average delay of requests in the queue for the $M/E_2/1$ system with delay. To do this, consider a system formed by two flows given by the functions of the distribution densities of the intervals:

– for the input flow

$$f_\lambda(t)=\lambda e^{-\lambda(t-t_0)}, \quad (20)$$

– for service times

$$f_\mu(t)=\mu^2(t-t_0)e^{-\mu(t-t_0)}. \quad (21)$$

We denote such a system $M^-/E_2^-/1$. Based on a similar statement 1, we conclude that for a pair of systems $M/E_2/1$ and $M^-/E_2^-/1$, their Laplace transforms, the delay density functions and formulas for the average delay of requests in the queue also coincide.

For the service time according to the law, we obtain similar expressions for the average service time and the coefficient of variation:

$$\bar{\tau}_\mu=2/\mu+t_0, \quad (22)$$

$$c_\mu=\sqrt{2}/(2+\mu t_0). \quad (23)$$

To determine the unknown parameters of distributions (20) and (21), we use the corresponding moment equations:

$$\bar{\tau}_\lambda=1/\lambda+t_0, \qquad \bar{\tau}_\mu=2/\mu+t_0, \qquad c_\lambda=(1+\lambda t_0)^{-1},$$

$$c_\mu=\sqrt{2}/(2+\mu t_0).$$

By specifying the values of the numerical characteristics and the shift parameter $t_0>0$ as input parameters for the system, and having determined the roots, we can calculate the average delay using formula (16).

4 EXPERIMENTS

Table 1–2 below shows the calculation data for systems $E_2^-/M^-/1$ and $M^-/E_2^-/1$ for cases of low, medium and high load.for cases of low, medium and high load $\rho=0.1;0.5;0.9$ for a wide range of c_λ, c_μ and a shift parameter t_0. For comparison, the right-hand columns show data for conventional systems $E_2/M/1$ and $M/E_2/1$.

Table 1 – Results of experiments for QS $E_2^-/M^-/1$ and $E_2/M/1$

Input parameters				Average delay	
ρ	c_λ	c_μ	t_0	For QS $E_2^-/M^-/1$	For QS $E_2/M/1$
0.1	0.643	0.1	0.9	0.000	0.030
	0.672	0.5	0.5	0.005	
	0.700	0.9	0.1	0.023	
	0.706	0.99	0.01	0.029	
0.5	0.389	0.1	0.9	0.003	0.618
	0.530	0.5	0.5	0.132	
	0.672	0.9	0.1	0.491	
	0.704	0.99	0.01	0.605	
0.9	0.134	0.1	0.9	0.055	6.588
	0.389	0.5	0.5	1.609	
	0.643	0.9	0.1	5.322	
	0.701	0.99	0.01	6.456	

Table 2 – Results of experiments for QS $M^-/E_2^-/1$ and $M/E_2/1$

Input parameters				Average delay	
ρ	c_λ	c_μ	t_0	For QS $M^-/E_2^-/1$	For QS $M/E_2/1$
0.1	0.643	0.071	0.9	0.001	0.083
	0.950	0.354	0.5	0.021	
	0.990	0.636	0.1	0.068	
	0.999	0.700	0.01	0.082	
0.5	0.550	0.071	0.9	0.008	0.75
	0.750	0.354	0.5	0.188	
	0.950	0.636	0.1	0.608	
	0.995	0.700	0.01	0.735	
0.9	0.190	0.071	0.9	0.068	6.75
	0.550	0.354	0.5	1.688	
	0.910	0.636	0.1	5.468	
	0.991	0.700	0.01	6.616	

DOI 10.15588/1607-3274-2022-1-8

Results for systems with a delay are compared with results for usual systems. It is obvious that the average waiting time in a system with a delay depends on the shift parameter t_0. The load factor ρ in both tables is determined by the ratio of average intervals $\rho = \overline{\tau}_\mu / \overline{\tau}_\lambda$. The calculations used the normalized service time $\overline{\tau}_\mu = 1$.

5 RESULTS

In this work, spectral expansions of the solution to the Lindley integral equation are obtained for the usual dual systems $E_2/M/1$ and $M/E_2/1$, as well as their analogs with shifted distribution laws, with the help of which the calculation formulas for the average delay of requests in the queue in a closed form are derived.

The same calculation formulas are valid for systems with time lag, respectively, taking into account changes in the numerical characteristics of their shifted distributions. The results of numerical calculations in Tables 1–2 are identical to the data obtained for the same systems with normalized Erlang distributions.

6 DISCUSSION

The average delay of requests in the queue in systems with latency is, as expected, less than in conventional systems, and as the value of the offset parameter decreases, it approaches the average waiting time in a conventional system. This fully confirms the adequacy of the constructed mathematical models.

Thus, the results of Table 1 and 2 confirm the complete adequacy of the constructed mathematical models for determining the average delay of requests in the queue both for ordinary dual systems and their analogs with shifted distribution laws.

In contrast to the conventional $E_2/M/1$ system, the $E_2^- / M^- / 1$ system with delay can be used for a range c_λ of 0 to $1/\sqrt{2}$ and c_μ 0 to 1. In the case of the $M/E_2/1$ system, the $M^- / E_2^- / 1$ system with delay allows a ramp range c_λ of 0 to 1, and c_μ from 0 to $1/\sqrt{2}$. Thus, the main advantage of introducing distributions shifted to the right from the zero point is to expand the range of variation coefficients of arrival intervals and service time.

Due to this, the scope of these QSs is expanding. Note that, in addition to the average delay of requests in the queue, it is possible to determine the variance and moments of higher orders of the delay time.

CONCLUSIONS

The problem of deriving formulas for the average delay of requests in the queue for two pairs of dual queuing systems with ordinary Erlang distributions in contrast to normalized distributions is solved.

The scientific novelty of the results is that spectral expansions of the solution of the Lindley integral equation for the systems under consideration are obtained and with their help the calculated formulas for the average delay in the queue for systems with delay in closed form are derived. These formulas complement and expands the well-known incomplete formula for the average waiting time in the G/G/1 systems with arbitrary laws of input flow distribution and service time.

The practical significance of the work lies in the fact that the obtained results can be successfully applied in the modern theory of teletraffic, where the delays of incoming traffic packets play a primary role. For this, it is necessary to know the numerical characteristics of the incoming traffic intervals and the service time at the level of the first two moments, which does not cause difficulties when using modern traffic analyzers.

Prospects for further research are seen in the continuation of the study of systems of type G/G/1 with other common input distributions and in expanding and supplementing the formulas for average waiting time.

ACKNOWLEDGEMENTS

This work was carried out as part of the author's scientific school "Methods and Models for the Research of Computing Systems and Networks", registered at the Russian Academy of Natural Sciences on 31.03.2015 and was supported by the University of PSUTI.

REFERENCES

1. Kleinrock L. Queueing Systems, Vol. I: Theory. New York: Wiley, 1975, 417 p.
2. Tarasov V. N. Extension of the Class of Queueing Systems with Delay, *Automation and Remote Control*, 2018, Vol. 79, No. 12, pp. 2147–2157. DOI: 10.1134/S0005117918120056
3. Tarasov V. N. Queuing systems with delay, *Radio Electronics, Computer Science, Control*, 2019, No. 3, pp. 55–63. DOI: 10.15588/1607-3274-2019-3-7
4. Tarasov V. N. The analysis of two queuing systems HE2/M/1 with ordinary and shifted input distributions, *Radio Electronics, Computer Science, Control*, 2019, No. 2, pp. 71–79. DOI: 10.15588/1607-3274-2019-2-8
5. Tarasov V. N. Analysis and comparison of two queueing systems with hypererlangian input distributions, *Radio Electronics, Computer Science, Control*, 2018, No.4, pp. 61–70. DOI: 10.15588/1607-3274-2018-4-6
6. Tarasov V. N., Bakhareva N. F. Comparative analysis of two queuing systems M/HE2/1 with ordinary and with the shifted input distributions, *Radio Electronics, Computer Science, Control*, 2019, No. 4, pp. 50–58. DOI: 10.15588/1607-3274-2019-4-5
7. Tarasov V. N. Analysis of H-2/E-2/1 system and her of the analog with shifted input distributions, *Radio Electronics, Computer Science, Control*, 2020, No. 1, pp. 90–97. DOI: 10.15588/1607-3274-2020-1-10
8. Do T. V., Chakka R., Sztrik J. Spectral Expansion Solution Methodology for QBD-M Processes and Applications in Future Internet Engineering, *ICCSAMA*, 2016, SCI 479, pp. 131–142. DOI: 10.1007/978-3-319-00293-4-11
9. Ma X., Wang Y., Zhu X., Liu W., Lan Q., Xiao W. A Spectral Method for Two-Dimensional Ocean Acoustic Propagation, *J. Mar. Sci. Eng.*, 2021, No. 9, pp. 1–19. DOI: https://doi.org/ 10.3390/jmse9080892
10. Brannstrom N. A. Queueing Theory analysis of wireless radio systems. Appllied to HS-DSCH. Lulea university of technology, 2004, 79 p.
11. Whitt W. Approximating a point process by a renewal process: two basic methods, *Operation Research*, 1982, Vol. 30, No. 1, pp. 125–147.

DOI 10.15588/1607-3274-2022-1-8

12. Myskja A. An improved heuristic approximation for the GI/GI/1 queue with bursty arrivals, *Teletraffic and datatraffic in a Period of Change, ITC-13. Elsevier Science Publishers*, 1991, pp. 683–688.
13. Aliev T. I. Approximation of Probability Distributions in Queuing Models, *Scientific and technical bulletin of information technologies, mechanics and optics*, 2013, No. 2, pp. 88–93.
14. Kruglikov V. K., Tarasov V. N. Analysis and calculation of queuing-networks using the two-dimensional diffusion-approximation, *Automation and Remote Control*, 1983, Vol. 44, No. 8. pp. 1026–1034.
15. Novitzky S., Pender J., Rand R.H., Wesson E. Limiting the oscillations in queues with delayed information through a novel type of delay announcement, *Queueing Systems*, 2020, Vol. 95, pp. 281–330. DOI: https://doi.org/10.1007/s11134-020-09657-9
16. Novitzky S., Pender J., Rand R.H., Wesson E. Nonlinear Dynamics in Queueing Theory: Determining the Size of Oscillations in Queues with Delay, *SIAM J. Appl. Dyn. Syst*., 18–1 2019, Vol. 18, No. 1, pp. 279–311. DOI: https://doi.org/10.1137/18M1170637
17. RFC 3393 [IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)] Available at: https://tools.ietf.org/html/rfc3393. (accessed: 26.02.2016).
18. Aras A. K., Chen X. & Liu Y. Many-server Gaussian limits for overloaded non-Markovian queues with customer abandonment, *Queueing Systems*, 2018, Vol. 89, No. 1, pp. 81–125. DOI: https://doi.org/10.1007/s11134-018-9575-0
19. Jennings O.B. & Pender J. Comparisons of ticket and standard queues, *Queueing Systems*, 2016, Vol. 84, No. 1, pp. 145–202. DOI: https://doi.org/10.1007/s11134-016-9493-y
20. Gromoll H.C., Terwilliger B. & Zwart B. Heavy traffic limit for a tandem queue with identical service times, *Queueing Systems*, 2018, Vol. 89, No. 3, pp. 213–241. DOI: https://doi.org/10.1007/s11134-017-9560-z
21. Legros B. M/G/1 queue with event-dependent arrival rates, *Queueing Systems*, 2018, Vol. 89, No. 3, pp. 269–301. DOI: https://doi.org/10.1007 /s11134-017-9557-7
22. Bazhba M., Blanchet J., Rhee CH., et al. Queue with heavy-tailed Weibull service times, *Queueing Systems*, 2019, Vol. 93, No. 11, pp. 1–32. DOI: https://doi.org/10.1007/s11134-019-09640-z/
23. Adan I., D'Auria B., Kella O. Special volume on 'Recent Developments in Queueing Theory' of the third ECQT conference, *Queueing Systems*, 2019, Vol. 93, No. 1, pp. 1–190. DOI: https://doi.org/10.1007/s11134-019-09630-1
24. Adan I., D'Auria B., Kella O. Special volume on 'Recent Developments in Queueing Theory' of the third ECQT conference: part 2, *Queueing Systems*, 2019, pp. 1–2. DOI: https://doi.org/10.1007/s11134-019-09637-8
25. Tibi D. Martingales and buffer overflow for the symmetric shortest queue model, *Queueing Systems,* Vol. 93, 2019, pp. 153–190. DOI: 10.1007/s11134-019-09628-9
26. Jacobovic R., Kella O. Asymptotic independence of regenerative processes with a special dependence structure, *Queueing Systems,* 2019, Vol. 93, pp. 139–152. DOI: 10.1007/s11134-019-09606-1
27. Wang L., Kulkarni V. Fluid and diffusion models for a system of taxis and customers with delayed matching, *Queueing Systems,* 2020, Vol. 96, pp. 101–131. DOI: 10.1007/s11134-020-09659-7

Received 25.11.2021.
Accepted 30.12.2021.

УДК 621.391.1:621.395

ДВІ ПАРИ ДВОЇСТИЙСИСТЕМ МАСОВОГО ОБСЛУГОВУВАННЯ ЗІ ЗВИЧАЙНИМИ І ЗСУНУТИМИ РОЗПОДІЛАМИ

Тарасов В. Н. – д-р техн. наук, професор, завідувач кафедри програмного забезпечення та управління в технічних системах Поволзького державного університету телекомунікацій та інформатики, Російська Федерація.

Бахарєва Н. Ф. – д-р техн. наук, професор, завідувач кафедри інформатики та обчислювальної техніки Поволзького державного університету телекомунікацій та інформатики, Російська Федерація.

АНОТАЦІЯ

Актуальність. Актуальність дослідження систем G/G/1 пов'язана з тим, що вони потрібні для моделювання систем передачі різного призначення, а також з тим, що для них не існує рішення в кінцевому вигляді в загальному випадку. Розглянуто задачу виведення рішення для середньої затримки вимог у черзі в замкнутій формі для звичайних систем з ерлангівським і експонентним вхідними розподілами і для цих систем зі зсунутими вправо розподілами.

Мета роботи. Отримання рішення для основної характеристики системи – середньої затримки вимог у черзі для двох пар систем масового обслуговування зі звичайними і зі зсунутими ерлангівськими та експоненціальними вхідними розподілами, а також порівняння результатів для систем із нормованими ерлангівськими розподілами. Отримання рішення для основної характеристики системи – середнього часу очікування вимог в черзі для двох систем масового обслуговування типу G/G/1 зі зсунутими вхідними розподілами.

Метод. Для вирішення поставленого завдання був використаний метод спектрального рішення інтегрального рівняння Ліндлі, який дозволяє отримати рішення для середньої затримки в черзі для розглянутих систем в замкнутій формі. Для практичного застосування отриманих результатів було використаний відомий метод моментів теорії ймовірностей.

Результати. Отримано спектральні рішення інтегрального рівняння Ліндлі для двох пар систем, за допомогою яких виведені розрахункові формули для середньої затримки вимог у черзі в замкнутій формі. Порівняння отриманих результатів зі даними для систем зі нормованими ерлангівськими розподілами підтверджує їхню ідентичність.

Висновки. Введення параметра зсуву в часі в закони розподілу вхідного потоку і часу обслуговування для систем, що розглядаються, перетворює їх в системи запізненням з меншим часом очікування. Це пов'язано з тим, що операція зсуву у часі зменшує величину коефіцієнтів варіацій інтервалів між надходженнями вимог та його часу обслуговування, а як відомо з теорії масового обслуговування, середня затримка вимог пов'язана з цими коефіцієнтами варіацій квадратичною

DOI 10.15588/1607-3274-2022-1-8

залежністю. Якщо система з ерлангівським і експонентним вхідними розподілами працює тільки при одній фіксованій парі значень коефіцієнтів варіацій інтервалів між надходженнями вимог та їх часу обслуговування, то ця ж система зі зрушеними розподілами дозволяє оперувати з інтервальними значеннями коефіцієнтів варіацій, що розширює сферу застосування цих систем. Аналогічно і зі зрушеними експонентними розподілами. Крім того, зрушений експонентний розподіл містить два параметри і дозволяє апроксимувати довільні закони розподілу з використанням перших двох моментів. Такий підхід дозволяє розрахувати середній час очікування та моменти вищих порядків для зазначених систем у математичних пакетах для широкого діапазону зміни параметрів трафіку. Метод спектрального вирішення інтегрального рівняння Ліндлі для розглянутих систем дозволив отримати рішення у замкнутій формі, і ці отримані рішення публікується вперше.

КЛЮЧОВІ СЛОВА: ерлангівський і експонентний закони розподілу, інтегральне рівняння Ліндлі, метод спектрального розкладання, перетворення Лапласа.

УДК 621.391.1:621.395

ДВЕ ПАРЫ ДВОЙСТВЕННЫХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С ОБЫЧНЫМИ И СДВИНУТЫМИ ЗАКОНАМИ РАСПРЕДЕЛЕНИЙ

Тарасов В. Н. – д-р техн. наук, профессор, заведующий кафедрой программного обеспечения и управления в технических системах Поволжского государственного университета телекоммуникаций и информатики, Российская Федерация.

Бахарева Н. Ф. – д-р техн. наук, профессор, заведующая кафедрой информатики и вычислительной техники Поволжского государственного университета телекоммуникаций и информатики, Российская Федерация.

АННОТАЦИЯ

Актуальность. Актуальность исследований систем G/G/1 связана с тем, что они востребованы для моделирования систем передачи данных различного назначения, а также с тем, что для них не существует решения в конечном виде в общем случае. Рассмотрена задача вывода решения для средней задержки требований в очереди в замкнутой форме для обычных систем с эрланговским и экспоненциальным входными распределениями и для этих же систем со сдвинутыми вправо распределениями.

Цель работы. Получение решения для основной характеристики системы – средней задержки требований в очереди для двух пар систем массового обслуживания с обычными и со сдвинутыми эрланговскими и экспоненциальными входными распределениями, а также сравнение результатов для систем с нормированными эрланговскими распределениями.

Метод. Для решения поставленной задачи использован метод спектрального решения интегрального уравнения Линдли, который позволяет получить решение для среднего времени ожидания для рассматриваемых систем в замкнутой форме. Для практического применения полученных результатов использован метод моментов теории вероятностей.

Результаты. Получены спектральные решения интегрального уравнения Линдли для двух пар систем, с помощью которых выведены расчетные формулы для средней задержки требований в очереди в замкнутой форме. Сравнение полученных результатов с данными для систем с нормированными эрланговскими распределениями подтверждает их идентичность.

Выводы. Введение параметра сдвига во времени в законы распределения входного потока и времени обслуживания для рассматриваемых систем, преобразует их в системы запаздыванием с меньшим временем ожидания. Это связано с тем, что операция сдвига во времени уменьшает величину коэффициентов вариаций интервалов между поступлениями требований и их времени обслуживания, а как известно из теории массового обслуживания, средняя задержка требований связана с этими коэффициентами вариаций квадратичной зависимостью. Если система с эрланговским и экспоненциальным входными распределениями работает только при одной фиксированной паре значений коэффициентов вариаций интервалов между поступлениями требований и их времени обслуживания, то эта же система со сдвинутыми распределениями позволяет оперировать с интервальными значениями коэффициентов вариаций, что расширяет область применения этих систем. Аналогично обстоит дело и со сдвинутыми экспоненциальными распределениями. Кроме того, сдвинутое экспоненциальное распределение содержит два параметра и позволяет аппроксимировать произвольные законы распределения с использованием двух первых моментов. Такой подход позволяет рассчитать среднее время ожидания и моменты высших порядков для указанных систем в математических пакетах для широкого диапазона изменения параметров трафика. Метод спектрального решения интегрального уравнения Линдли для рассматриваемых систем позволил получить решение в замкнутой форме и эти полученные решения публикуется впервые.

КЛЮЧЕВЫЕ СЛОВА: эрланговский и экспоненциальный законы распределения, интегральное уравнение Линдли, метод спектрального разложения, преобразование Лапласа.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Kleinrock L. Queueing Systems, Vol. I: Theory / L. Kleinrock – New York: Wiley, 1975. – 417 p.
2. Tarasov V. N. Extension of the Class of Queueing Systems with Delay / V. N. Tarasov // Automation and Remote Control. – 2018. – Vol. 79. – No. 12. – P. 2147–2157. DOI: 10.1134/S0005117918120056
3. Tarasov V. N. Queueing systems with delay / V. N. Tarasov // Radio Electronics, Computer Science, Control. – 2019. – No. 3. P. 55–63. DOI: 10.15588/1607-3274-2019-3-7
4. Tarasov V. N. The analysis of two queuing systems HE2/M/1 with ordinary and shifted input distributions / V. N. Tarasov // Radio Electronics, Computer Science, Control. – 2019. – No. 2. – P. 71–79. DOI: 10.15588/1607-3274-2019-2-8
5. Tarasov V.N. Analysis and comparison of two queueing systems with hypererlangian input distributions /

DOI 10.15588/1607-3274-2022-1-8

V. N. Tarasov // Radio Electronics, Computer Science, Control. – 2018. – No. 4. – P. 61–70. DOI: 10.15588/1607-3274-2018-4-6
6. Tarasov V. N. Comparative analysis of two queuing systems M/HE2/1 with ordinary and with the shifted input distributions / V. N. Tarasov, N. F. Bakhareva // Radio Electronics, Computer Science, Control. – 2019. – No. 4. – P. 50–58. DOI: 10.15588/1607-3274-2019-4-5
7. Tarasov V. N. Analysis of H-2/E-2/1 system and her of the analog with shifted input distributions / V. N. Tarasov // Radio Electronics, Computer Science, Control. – 2020. – No. 1. P. 90–97. DOI: 10.15588/1607-3274-2020-1-10
8. Do T. V. Spectral Expansion Solution Methodology for QBD-M Processes and Applications in Future Internet Engineering / T. V. Do, R. Chakka, J. Sztrik // ICCSAMA. – 2016. – SCI 479. P. 131–142. DOI: 10.1007/978-3-319-00293-4-11
9. Ma X. A Spectral Method for Two-Dimensional Ocean Acoustic Propagation / X. Ma, Y. Wang, X. Zhu, W. Liu, Q. Lan, W. Xiao // J. Mar. Sci. Eng. – 2021. – No. 9. P. 1–19. DOI: https://doi.org/ 10.3390/jmse9080892
10. Brannstrom N. A Queueing Theory analysis of wireless radio systems / N. Brannstrom – Appllied to HS-DSCH. Lulea university of technology, 2004. –79 p.
11. Whitt W. Approximating a point process by a renewal process: two basic methods / W. Whitt // Operation Research. – 1982. – № 1. – P. 125–147.
12. Myskja A. An improved heuristic approximation for the GI/GI/1 queue with bursty arrivals / A. Myskja // Teletraffic and datatraffic in a Period of Change, ITC-13. Elsevier Science Publishers. – 1991. – P.683–688.
13. Aliev T. I. Approximation of Probability Distributions in Queuing Models / T. I. Aliev // Scientific and technical bulletin of information technologies, mechanics and optics. – 2013. – No. 2. – P. 88–93.
14. Kruglikov V. K. Analysis and calculation of queuing-networks using the two-dimensional diffusion-approximation / V. K. Kruglikov, V. N. Tarasov // Automation and Remote Control. – 1983. – Vol. 44, № 8. – P. 1026–1034.
15. Limiting the oscillations in queues with delayed information through a novel type of delay announcement / [S. Novitzky, J. Pender, R. H. Rand, E. Wesson] // Queueing Systems. – 2020. – P. 281–330. DOI: https://doi.org/10.1007/s11134-020-09657-9
16. Novitzky S. Nonlinear Dynamics in Queueing Theory: Determining the Size of Oscillations in Queues with Delay / S. Novitzky, J. Pender, R.H. Rand, E. Wesson // SIAM J. Appl. Dyn. Syst. – 2019. – № 1. – P. 279–311. DOI: https://doi.org/10.1137/18M1170637
17. HTTPS://tools.ietf.org/html/rfc3393. RFC 3393 IP Packet Delay Variation Metric for IP Performance Metrics (IPPM) (дата обращения: 26.02.2016).
18. Aras A. K. Many-server Gaussian limits for overloaded non-Markovian queues with customer abandonment / A. K. Aras, X. Chen, Y. Liu // Queueing Systems. – 2018. – No. 1. – P. 81–125. DOI: https://doi.org/10.1007/s11134-018-9575-0
19. Jennings O. B. Comparisons of ticket and standard queues / O. B. Jennings, J. Pender // Queueing Systems. – 2016. – No. 1. – P. 145–202. DOI: https://doi.org/10.1007/s11134-016-9493-y
20. Gromoll H. C. Heavy traffic limit for a tandem queue with identical service times / H. C. Gromoll, B. Terwilliger, B. Zwart // Queueing Systems. – 2018. – No. 3. – P. 213–241. DOI: https://doi.org/10.1007/s11134-017-9560-z
21. Legros B. M/G/1 queue with event-dependent arrival rates / B. Legros // Queueing Systems. – 2018. – No. 3. – P. 269–301. DOI: https://doi.org/10.1007 /s11134-017-9557-7
22. Bazhba M. Queue with heavy-tailed Weibull service times / M. Bazhba, J. Blanchet, C. H. Rhee // Queueing Systems. – 2019. – No. 11. – P. 1–32. DOI: https://doi.org/10.1007/s11134-019-09640-z/
23. Adan I. Special volume on ‘Recent Developments in Queueing Theory’ of the third ECQT conference / I. Adan, B. D’Auria, O. Kella // Queueing Systems. – 2019. – No. 1. – P. 1–190. DOI: https://doi.org/10.1007/s11134-019-09630-1
24. Adan I. Special volume on ‘Recent Developments in Queueing Theory’ of the third ECQT conference: part 2 / I. Adan, B. D’Auria, O. Kella // Queueing Systems. – 2019. – P. 1–2. DOI: https://doi.org/10.1007/s11134-019-09637-8
25. Tibi D. Martingales and buffer overflow for the symmetric shortest queue model / D. Tibi // Queueing Systems. – 2019. – P. 153–190. DOI: 10.1007/s11134-019-09628-9
26. Jacobovic R. Asymptotic independence of regenerative processes with a special dependence structure / R. Jacobovic, O. Kella // Queueing Systems. – 2019. – P. 139–152. DOI: 10.1007/s11134-019-09606-1
27. Wang L. Fluid and diffusion models for a system of taxis and customers with delayed matching / L. Wang, V. Kulkarni // Queueing Systems. – 2020. – P. 101–131. DOI: 10.1007/s11134-020-09659-7

DOI 10.15588/1607-3274-2022-1-8

e-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2022. № 1
p-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2022. № 1

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

УДК 004.8:004.032.26

ШВИДКА НЕЧІТКА ПРАВДОПОДІБНА КЛАСТЕРИЗАЦІЯ НА ОСНОВІ АНАЛІЗУ ПІКІВ ЩІЛЬНОСТІ РОЗПОДІЛУ ДАНИХ

Бодянський Є. В. – д-р техн. наук, професор, професор кафедри штучного інтелекту, Харківський національний університет радіоелектроніки, Харків, Україна.

Плісс І. П. – канд. техн. наук, провідний науковий співробітник ПНДЛ АСУ, Харківський національний університет радіоелектроніки, Харків, Україна.

Шафроненко А. Ю. – канд. техн. наук, доцент, доцент кафедри інформатики, Харківський національний університет радіоелектроніки, Харків, Україна.

АНОТАЦІЯ

Актуальність. Проблема кластеризації (класифікації без вчителя), що часто зустрічається при обробці масивів даних різної природи, є досить цікавою і невід'ємною частиною штучного інтелекту. Для вирішення цього завдання існує безліч відомих методів та алгоритмів, які базуються на принципах щільності розподілу спостережень в даних, що аналізуються. Однак ці методи досить складні в програмній реалізації та не позбавлені недоліків, а саме: проблеми визначення значущих кластерів в наборах даних різної щільності, багатоепохове самонавчання, застрягання в локальних екстремумах цільових функцій, тощо. Слід зазначити, що методи, засновані на аналізі піків щільності розподілу даних, є за своєю природою чіткими, тому для розширення можливостей цих методів доцільно ввести їх нечітку модифікацію.

Мета. Мета роботи полягає у запровадженні швидкої нечіткої кластеризації даних з використанням піків щільності розподілу даних, яка може знаходити екстемуми (центоїди) кластерів, що перетинаються незалежно від кількості даних, що надходять.

Метод. Розглянуто задачу нечіткої кластеризації масивів даних на основі гібридного методу, заснованого на одночасному використанні правдоподібного підходу до нечіткої кластеризації і алгоритму знаходження типів щільності розподілу вихідних даних. Особливістю запропонованого методу є обчислювальна простота і висока швидкість, пов'язана з тим, що весь масив обробляється тільки один раз, тобто виключається необхідність в багатоепоховому самонавчанні, що реалізується в традиційних алгоритмах нечіткої кластеризації.

Результати. Особливістю запропонованого методу швидкої нечіткої правдоподібної кластеризації на основі аналізу піків щільності розподілу даних є обчислювальна простота і висока швидкість, пов'язана з тим, що весь масив обробляється тільки один раз, тобто виключається необхідність у багатоепоховому самонавчанні, що реалізується в традиційних алгоритмах нечіткої кластеризації. Результати обчислювального експерименту підтверджують ефективність запропонованого підходу в задачах кластерзації в умовах, коли кластери перетинаються.

Висновки. Результати експерименту дозволяють рекомендувати розроблений метод для вирішення проблем автоматичної кластеризації та класифікації даних та максимально швидко знаходити центри кластерів. Запропонований метод швидкої нечіткої правдоподібної кластеризації на основі аналізу піків щільності розподілу даних призначений для використання в системах обчислювального інтелекту, нейро-фаззі системах,в навчанні штучних нейронних мереж та у завданнях кластеризації.

КЛЮЧОВІ СЛОВА: нечітка кластеризація, правдоподібна кластеризація, піки щільності розподілу даних.

АБРЕВІАТУРА

DBSCAN density-based spatial clustering of applications with noise;

OPTICS ordering points to identify the clustering structure;

NMI нормалізована взаємна інформація;

FCDP метод швидкої нечіткої правдоподібної кластеризації на основі аналізу піків щільності розподілу.

НОМЕНКЛАТУРА

X – матриця набору даних;

k – номер вектору-спостереження;

DOI 10.15588/1607-3274-2022-1-9

i – номер атрибуту вектора-спостереження;

j – номер класу;

$x(k)$ – вектор-спостереження;

l, q – номери кластерів;

m – кількість неперетинних класів;

μ_j – рівень нечіткої належності j-го кластеру;

D – матриця відстаней між спостереженнями;

d – відстань між спостереженнями;

ρ – вектор локальної щільності;

c – центроїд кластера;

δ_k^* – точка з максимальною щільністю;

Cr – рівень правдоподібності.

ВСТУП

Задача кластеризації (класифікації без вчителя) часто зустрічається при обробці масивів спостережень самої різної природи в рамках загальної проблеми Data Mining, Big Data Mining, Data Stream Mining, тощо. Для вирішення цієї задачі, існує безліч підходів від найпростіших (*k*-середніх, ієрархічних) алгоритмів до найбільш просунутих, заснованих на аналізі щільності розподілу даних

Слід зазначити, що методи, засновані на аналізі піків щільності розподілу даних, є за своєю природою чіткими, тому для розширення можливостей цих методів доцільно ввести їх нечітку модифікацію.

Об'єкт дослідження швидка нечітка кластеризація даних на основі піків щільності розподілу даних.

Предмет дослідження процедура аналізу піків щільності розподілу даних.

Мета роботи полягає у запровадженні швидкої нечіткої кластеризації даних з використанням піків щільності розподілу даних, яка може знаходити екстемуми (центри) кластерів, що перетинаються незалежно від кількості даних, що надходять.

1 ПОСТАНОВКА ЗАВДАННЯ

Процес нечіткої кластеризації на основі аналізу піків щільності розподілу даних зручно представити у вигляді послідовності кроків, при цьому вихідною інформацією як і в інших методах, заснованих на парадигмі самонавчання, є нерозмічена вибірка векторних спостережень $X = \{x(1), x(2), ..., x(k), ..., x(N)\}$, $x(k) = \{x_i(k)\} \in R^n$, при цьому для зручності розрахунків всі компоненти цих векторів попередньо закодовані в деякому обмеженому інтервалі, наприклад, $-1 \le x_i(k) \le 1 \forall i, k$.

2 ОГЛЯД ЛІТЕРАТУРИ

Для вирішення цієї задачі існує безліч підходів[1] від найпростіших типу *k*-середніх і ієрархічних алгоритмів до найбільш просунутих, заснованих на аналізі щільності розподілу даних [2–4]. Найбільш популярними з таких алгоритмів є DBSCAN [5], DENCLUE [6], OPTICS [7] та їм подібні, що дозволяють оцінити не тільки центроїди, але і їх кількість. У той же час алгоритми, які засновані на щільностях, досить складні з точки зору чисельної реалізації, а процедури градієнтної оптимізації, що лежать в їх основі, схильні до застрягання в локальних екстремумах цільових функцій.

Цих недоліків позбавлений метод, заснований на пошуку піків щільності розподілу даних [8]. Слід зазначити, що більш відомі алгоритми, які аналізують щільність розподілу даних, не виконують пошук екстремумів, а призначені для вирішення завдань чіткої кластеризації, тобто в умовах коли апріорно відомо, що кластери, які формуються, взаємно не перетинаються, а кожне спостереження з вихідного масиву може належати тільки одному кластеру.

У той же час в реальних задачах досить часто зустрічається задачі, коли будь-яке з спостережень із аналізованого масиву даних, може з різними рівнями належності одночасно належати до декількох класів-кластерів. Ця ситуація розглядається в рамках нечіткого кластерного аналізу [9], при цьому історично склалося два підходи до вирішення цієї проблеми: імовірнісний і можливісний.

Зауважимо також, що в останні роки з'явився, так званий, довірчий підхід [10, 11], що володіє перевагами як імовірнісного, так і можливісного підходів.

3 МАТЕРІАЛИ І МЕТОДИ

В процесі кластеризації на основі аналізу піків щільності розподілу даних аналізується два параметри: ρ_k – локальна щільність і δ_k – відстань до точки з більш високою щільністю. Окрім того вводиться єдиний вільний параметр d_c – відстань зрізу, яка задається і варіюється користувачем для отримання необхідної точності рішення задачі.

Роботу методу можна сформулювати як наступну послідовність елементарних кроків:

1. На першому кроці на основі вихідної $(n \times N)$ матриці «об'єкт – властивість» вводиться $(N \times N)$-матриця відстаней між спостереженнями:

$$D = \{d_{kl}\},\ d_{kl} = \|x(k) - x(l)\| \forall k, l,$$

при цьому може бути використана будь-яка метрика, яка використовується в Data Mining і, зокрема, в кластерному аналізі.

2. На другому кроці розраховується $(N \times 1)$-вектор локальних щільностей $\rho = \{\rho_k\} \in R^N$:

$$\rho_k = \sum_{k=1}^{N} \chi(d_{kl} - d_c),$$

DOI 10.15588/1607-3274-2022-1-9

де $\chi(d) = \begin{cases} 1, \text{ якщо } d < 0, \\ 0, \text{ в іншому випадку.} \end{cases}$

Відстань зрізу обирається з суто емпіричних міркувань, при цьому автори методу [8] рекомендують вибирати його так, щоб в околі, який формується, потрапляло $0,01N - 0,02N$ спостережень вибірки, що оброблюється.

3. Розрахунок вектора мінімальних відстаней $\delta = \{\delta_k\} \in R^N$ до точок з більш високою щільністю

$$\delta_k = \min_{\forall l, \rho_l > \rho_k} \{d_{kl}\},$$

а для точки з максимальною щільністю δ_k^* розраховується:

$$\delta_k^* = \max_l \{d_{kl}\}.$$

4. Формування центроїдів кластерів $c_j, j = 1,2,...,m,$ при цьому в якості центроїдів $c_j = x(k)$ обираються точки з найвищою щільністю, тобто обираються деякі спостереження з вихідної вибірки X. До кожного з центроїдів c_j приписуються точки, найближчі до нього в сенсі

$$\min(d_{kl}) \equiv d_{jl}.$$

Зауважимо також, що в [12] в якості центроїдів пропонується використовувати значення $c_j = x(k)$ з максимальним значенням добутків $\rho_k \cdot \delta_k$.

Далі всі центроїди впорядковуються за зменшенням цього добутку $c_1,...,c_j,...c_m$, а якість одержуваного рішення оцінюється за допомогою будь-якого з критеріїв, прийнятих в чіткій кластеризації [1].

Якщо з точки зору використаного критерію якість кластеризації виявляється незадовільною, можна або зменшити значення d_c, або збільшити число можливих кластерів, тобто $j = 1,...,m, m+1, m+2,....$

Далі процедура нечіткої кластеризації повторюється, починаючи з першого кроку.

5. Починаючи з п'ятого кроку реалізується процедура нечіткої кластеризації. При цьому для кожної точки $x(k) \neq c_j$ розраховуються рівні нечіткої належності в стандартній формі [9]

$$\mu_j(k) = \frac{d_{jk}^{-2}}{\sum_{l=1}^{m} d_{lk}^{-2}}, \quad (1)$$

або на основі функції щільності розподілу Коші [13]

$$\mu_j(k) = \left(1 + \frac{d_{jk}^{-2}}{\sigma_j^2}\right), \quad (2)$$

де

$$\sigma_j^2 = \left(\sum_{\substack{l=1 \\ l \neq k}}^{m} d_{lk}^{-2}\right)^{-1}.$$

6. На основі оцінок ймовірнісної нечіткої належності (1), (2) розраховується рівень довіри до отриманих результатів на основі стандартного правдоподібного підходу [10, 11]

$$Cr_j(k) = \frac{1}{2}\left(\mu_j^*(k) + 1 - \sup \mu_j^*(k)\right), \quad (3)$$

де

$$\mu_j^*(k) = \frac{\mu_j(k)}{\sup \mu_l(k)}.$$

7. Завершення процедури нечіткої кластеризації шляхом оцінки якості результатів за допомогою будь-якого з критеріїв, що застосовуються в нечіткої кластеризації [9], хоча оцінка (3) вже сама по собі надає наскільки можна довіряти правдоподібності отриманих результатів.

4 ЕКСПЕРИМЕНТИ

Експериментальні дослідження методу швидкої нечіткої правдоподібної кластеризації на основі аналізу піків щільності розподілу(FCDP) даних був реалізований на трьох масивах даних: табл. 1. Порівняльний аналіз проведено з відомими методами кластеризації які використовують параметр піків щільності розподілу даних, а саме: *k*-середніх, DBSCAN, який для заданої множини точок у деякому просторі відносить в одну групу точки, які розташовані найбільш щільно та розмічає точки, які лежать в областях з невеликою щільністю; DENCLUE, в якому кластери визначаються локальними максимумами оцінки щільності; OPTICS – алгоритм знаходження щільності на основі кластерів у просторових даних, що вирішує проблему визначення значущих кластерів в наборах даних різної щільності. За допомогою цих алгоритмів було проведено аналіз якості кластеризації на основі цих вибірок. Заздалегідь, для порівняльного аналізу із вибірок бралась частина спостережень і проводився аналіз якості кластеризації даних, яка виміряна показником нормалізованої взаємної інформації (приймає значення 1, якщо ідеальна кластеризація даних знайдена).

DOI 10.15588/1607-3274-2022-1-9

Таблиця 1 – Зразки даних

Назва вибірки	Кількість спостережень	Кількість атрибутів
iris	150	4
wine	178	13
ecoli	336	8

5 РЕЗУЛЬТАТИ

За результатами аналізу кластеризації була отримана інформація, яка представлена у табл. 2. Для кожної з вибірок перевірили, як розмір вибірки впливає на якість кластеризації.

Таблиця 2 – Значення показника нормалізованої взаємної інформації для різних даних та методів, перше число у трьох правих стовпцях показує розмір вибірки

Data		k-means	FCDP	DBSCAN	DENCLUE	OPTICS
iris	0,8	0,67±0,06	0,79±0,03	0,68±0,06	0,67±0,06	0,78±0,02
	0,4	0,65±0,07	0,68±0,18	0,60±0,06	0,67±0,06	0,72±0,08
	0,2	0,64±0,07	0,64±0,10	0,54±0,04	0,54±0,07	0,64±0,06
wine	0,8	0,70±0,11	0,78±0,02	0,66±0,01	0,68±0,01	0,76±0,04
	0,4	0,70±0,05	0,78±0,04	0,62±0,00	0,72±0,00	0,72±0,08
	0,2	0,58±0,21	0,69±0,11	0,48±0,01	0,70±0,01	0,59±0,01
ecoli	0,8	0,65±0,02	0,77±0,09	0,75±0,07	0,75±0,11	0,75±0,05
	0,4	0,65±0,04	0,75±0,05	0,63±0,01	0,73±0,05	0,73±0,07
	0,2	0,65±0,03	0,70±0,10	0,55±0,01	0,66±0,21	0,68±0,11

На рис. 1 продемонстрована залежність нормалізованої взаємної інформації (NMI) від розміру навчальної вибірки, що дає змогу говорити про те, що розмір вибірки не впливає на якість кластеризації, а NMI не є лінійним. Таким чином, якість кластеризації не втрачається навіть при 20% наявності вибірки.

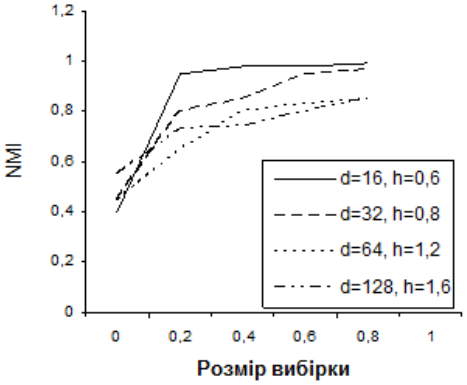


Рисунок 1 – Залежність показника нормалізованої взаємної інформації (NMI) від розміру навчальної вибірки

6 ОБГОВОРЕННЯ

За результатами експериментальних досліджень та аналізу отриманих результатів, можна зробити висновок, що запропонований метод швидкої нечіткої правдоподібної кластеризації на основі аналізу піків щільності розподілу даних порівняно з методами, заснованими на використанні піків щільності, демонструє гарні результати роботи.

Порівняльний аналіз запропонованого методу продемонстровано в табл. 2, в якій наведені значення показника нормалізованої взаємної інформації для різних даних та методів, перше число у трьох правих стовпцях показує розмір вибірки. Аналізуючи табл. 2, можна зробити висновки, що якість кластеризації даних не втрачається від кількості наявних спостережень у вибірці, тобто, незалежно від 20%, 40% або 80% наявності вибірки, якість кластеризації не зменшується.

Як видно із порівняльної табл. 2, показник нормалізованої взаємної інформації (NMI), що приймає значення 1, якщо ідеальна кластеризація даних знайдена серед всіх запропонованих методів кластеризацій найкращій результат демонструє, коли даних все-ж таки більше. Якщо порівнювати якість кластеризації даних з відомими методами, можна зробити висновок, що запропонований метод швидкої нечіткої правдоподібної кластеризації на основі аналізу піків щільності розподілу даних (FCDP) демонструє значно вищі показники ніж *k*-means, DBSCAN і DENCLUE та майже однаково з методом OPTICS. Так, якщо більш детально проаналізувати результат роботи цих двох методів по кількості спостережень, показник нормалізованої взаємної інформації у методі FCDP трошки вищій за OPTICS незалежно від виду вибірки, що подається на кластеризацію.

ВИСНОВКИ

Розглянуто задачу нечіткої кластеризації масивів даних на основі гібридного методу, заснованого на одночасному використанні правдоподібного підходу до нечіткої кластеризації і алгоритму знаходження типів щільності розподілу вихідних даних. Особливістю запропонованого методу є обчислювальна простота і висока швидкість, пов'язана з тим, що весь масив обробляється тільки один раз, тобто виключається необхідність в багатоепоховому самонавчанні, що реалізується в традиційних алгоритмах нечіткої кластеризації. Результати обчислювального експерименту підтверджують ефективність запропонованого підходу в задачах кластерзації в умовах коли кластери перетинаються.

Наукова новизна: вперше запропонована процедура швидкої нечіткої кластеризації даних з використанням піків щільності розподілу даних на основі правдоподібного підходу.

Практичне значення: результати експерименту дозволяють рекомендувати запропоновані методи для використання на практиці для вирішення проблем автоматичної кластеризації великих даних.

Перспективи подальших досліджень методи нечіткої кластеризації даних для широкого класу практичних проблем.

DOI 10.15588/1607-3274-2022-1-9

ПОДЯКА

Робота виконана в рамках науково-дослідного проєкту державного бюджету Харківського національного університету радіоелектроніки «Глибокі гібридні системи обчислювального інтелекту для аналізу потоків даних та їх швидке навчання» (номер державної реєстрації 0119U001403).

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Xu R. Clustering / R. Xu, D. C. Wunsch. – Hoboken N. J.: John Wiley & Sons, Inc., 2009. – 398 p.
2. Nadaraya E. A. On nonparametric estimates of density function and regression curves / E. A. Nadaraya // Theory of Probabilistic Application. – 1965. – № 10 – P. 186–190.
3. Epanechnikov V. A. Nonparametric estimation of multivariate probability density / V. A. Epanechnikov // Probability theory and its Application. – 1968. – 14, №2. – P. 156–161.
4. Fukunaga K. The estimation of the gradient of a density function with application in pattern recognition / K. Fukunaga, L. D. Hostler // IEEE Trans. on Inf. Theory, Jan., 1975. – IEEE. – 1975. – № 21 – P. 32–40. DOI: 10.1109/TIT.1975.10 55330.
5. Ester M. A density – based algorithm for discovering clusters in large spatial databases with noise / [M. Ester, H. Kriegel, J. Sandler, X. Xu] // Proc. 2nd Int. Conf. on Knowledge Discovering and Data Mining. – KDD96, N.Y.: AAAI Press, Aug. 2, 1996. – P. 226–231.
6. Hinneburg A. An efficient approach to clustering in large multimedia databases with noise / A. Hinneburg, D. Klein // Proc. 4th Int. Conf. in Knowledge Discovering and Data Mining. – KDD98, N.Y.: AAAI Press, Aug. 27, 1998. – Hinneburg, 1998. – P. 58–65.
7. OPTICS: Ordering points to identify the clustering structure / [M. Ankerst, M. Brening, H. Kriegel, J. Sander] // Proc. 1999 ACM SIGMOD Int. Conf. on Management of Data, Jun. 1, 1999. – Philadelphia, 1999. – P. 49–60.
8. Rodriguez A. Clustering by fast seach and find of density peaks / A. Rodriguez, A. Laio. – Science. – 2014. – № 34. – P. 1492–1496.
9. Fuzzy Clustering Analysis: Methods for Classification, Data Analisys and Image Recognition / [F. Höppner, F. Klawonn, R. Kruse, T. Runkler]. – Chichester : John Wiley &Sons, 1999. – 300 p.
10. Credibilistic clustering: the model and algorithms / [J. Zhou, Q. Wang, C.-C. Hung, X. Yi] // International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. – 2015. – Vol. 23, № 4. – P. 545–564. DOI: https://doi.org/10.1142/S0218488515500245
11. Zhou J. Credibilistic clustering algorithms via alternating cluster estimation / J. Zhou, Q. Wang, C. C. Hung // Journal of Intelligent Manufacturing. – 2017. – Vol. 28. – P. 727–738. DOI: https://doi.org/10.1007/s10845-014-1004-6.
12. Begum N. Accelerating dynamic time warping clustering with a novel admissible pruning strategy / [N. Begum, L. Ulanova, J. Wamg, E. Klogh] // Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 10, 2015. – Sydney, NSW, Australia. – P. 49–58. DOI: https://doi.org/10.1145/2783258.27 83286.
13. Online credibilistic fuzzy clustering of data using membership functions of special type [Electronic resource] / [A. Shafronenko, Ye. Bodyanskiy, I. Klymova, O. Holovin] // Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), April 27–1 May 2020. – Zaporizhzhia, 2020. – Access mode: http://ceur-ws.org/Vol-2608/paper56.pdf.

Стаття надійшла до редакції 28.10.2021.
Після доробки 25.12.2021.

УДК 004.8:004.032.26

БЫСТРАЯ НЕЧЕТКАЯ ПРАВДОПОДОБНАЯ КЛАСТЕРИЗАЦИЯ НА ОСНОВЕ АНАЛИЗА ПИКОВ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ ДАННЫХ

Бодянский Е. В. – д-р техн. наук, профессор, профессор кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники, Харьков, Украина.

Плісс І. П. – канд. техн. наук, ведущий научный сотрудник ПНДЛ АСУ Харьковского национального университета радиоэлектроники, Харьков, Украина.

Шафроненко А. Ю. – канд. техн. наук, доцент, доцент кафедры информатики Харьковского национального университета радиоэлектроники, Харьков, Украина.

АННОТАЦИЯ

Актуальность. Проблема кластеризации (классификации без учителя) часто встречается при обработке массивов данных различной природы и является достаточно интересной и неотъемлемой частью искусственного интеллекта. Для решения этой задачи существует множество известных методов и алгоритмов основанных на анализе плотности распределения наблюдений в анализируемых данных. Однако эти методы достаточно сложны в программной реализации и не лишены недостатков, а именно: проблемы определения значимых кластеров в наборах данных различной плотности, многоэпоховое самообучение, застревание в локальных экстремумах целевых функций и тому подобное. Следует отметить, что методы, основанные на анализе пиков плотности распределения данных, являются по своей природе четкими, поэтому для расширения возможностей этих методов целесообразно ввести их нечеткую модификацию.

Цель. Цель работы заключается в введении быстрой процедуре нечеткой кластеризации данных с использованием пиков плотности распределения данных, которая может находить экстемумы (центры) кластеров, которые пересекаются независимо от количества поступающих данных.

Метод. Рассмотрена задача нечеткой кластеризации массивов данных на основе гибридного метода, основанного на одновременном использовании правдоподобного подхода к нечеткой кластеризации и алгоритма нахождения типов плотности распределения исходных данных. Особенностью предлагаемого метода является вычислительная простота и высокая скорость, связанная с тем, что весь массив обрабатывается только один раз, то есть исключается необходимость в многоэпоховом самообучении, реализуемом в традиционных алгоритмах нечеткой кластеризации.

Результаты. Особенностью предложенного метода быстрой нечеткой правдоподобной кластеризации на основе анализа пиков плотности распределения данных является вычислительная простота и высокая скорость, связанная с тем, что весь массив обрабатывается только один раз, то есть исключается необходимость в многоэпоховом самообучении, что реализуется в традиционных алгоритмах нечеткой кластеризации. Результаты вычислительного эксперимента подтверждают эффективность предложенного подхода в задачах кластеризации в условиях, когда кластеры пересекаются.

Выводы. Результаты эксперимента позволяют рекомендовать разработанный метод для решения проблем автоматической кластеризации и классификации данных, максимально быстро находить центры кластеров. Предложенный метод бы-

DOI 10.15588/1607-3274-2022-1-9

строй нечеткой правдоподобной кластеризации на основе анализа пиков плотности распределения данных предназначен для использования в системах вычислительного интеллекта, нейро-фаззи системах, в обучении искусственных нейронных сетей и в задачах кластеризации.

КЛЮЧЕВЫЕ СЛОВА: нечеткая кластеризация, правдоподобная кластеризация, пики плотности распределения данных.

UDC 004.8:004.032.26

FAST FUZZY CREDIBILISTIC CLUSTERING BASED ON DENSITY PEAKS DISTRIBUTION OF DATA BROAKYSIS

Bodyanskiy Ye. V. – Dr. Sc., Professor at the Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Pliss I. P. – PhD, Leading Researcher at Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Shafronenko A. Yu. – PhD, Associated Professor at the Department of Informatics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

ABSTRACT

Context. The problem of clustering (classification without a teacher) is often occures when processing data arrays of various natures, which is quite an interesting and integral part of artificial intelligence. To solve this problem, there are many known methods and algorithms based on the principles of the distribution density of observations in the analyzed data. However, these methods are rather complicated in software implementation and are not without drawbacks, namely: the problem of determining significant clusters in datasets of different densities, multiepoch self-learning, getting stuck in local extrema of goal functions, etc. It should be noted that the methods based on the analysis of the peaks of the data distribution density are clear in nature, therefore, to expand the capabilities of these methods, it is advisable to introduce their fuzzy modification.

Objective. The aim of the work is to introduce fast fuzzy data clustering using density peaks distribution of the datasets, that can find the prototypes (centroids) of clusters that overlapping regardless of the amount of incoming data.

Method. The problem of fuzzy clustering data arrays based on a hybrid method that based on the simultaneous use of a credibilistic approach to fuzzy clustering and an algorithm for finding the types of distribution density of the initial data is proposed. A feature of the proposed method is computational simplicity and high speed, due to the fact that the entire array is processed only once, that is, eliminates the need for multi-era self-learning, implemented in traditional fuzzy clustering algorithms.

Results. A feature of the proposed method of fast fuzzy credibilistic clustering using of density peaks distribution is characterized by computational simplicity and high speed due to the fact that the entire array is processed only once, that is, the need for multiepoch self-learning is eliminated, which is implemented in traditional fuzzy clustering algorithms. The results of the computational experiment confirm the effectiveness of the proposed approach in clustering problems under conditions in the case when the clusters are ovelap.

Conclusions. The experimental results allow us to recommend the developed method for solving the problems of automatic clustering and data classification, as quickly as possible to find the centroids of clusters. The proposed method of fast fuzzy credibilistic clustering using of density peaks distribution of dataset is intended for use in computational intelligence systems, neuro-fuzzy systems, in training artificial neural networks and in clustering problems.

KEYWORDS: fuzzy clustering, credibilistic clustering, density peak of dataset.

REFERENCES

1. Xu R., Wunsch D. C. Clustering. Hoboken N.J., John Wiley & Sons, Inc., 2009, 398 p.
2. Nadaraya E. A. On nonparametric estimates of density function and regression curves, *Theory of Probabilistic Application*, 1965, No. 10, pp. 186–190.
3. Epanechnikov V. A. Nonparametric estimation of multivariate probability density, *Probability theory and its Application*, 1968, 14, No. 2, pp. 156–161.
4. Fukunaga K., Hostler L. D. The estimation of the gradient of a density function with application in pattern recognition, *IEEE Trans. on Inf. Theory, Jan., 1975, IEEE*, 1975, No. 21, pp. 32–40. DOI: 10.1109/TIT.1975.10 55330.
5. Ester M., Kriegel H., Sandler J., Xu X. A density – based algorithm for discovering clusters in large spatial databases with noise, *Proc. 2nd Int. Conf. on Knowledge Discovering and Data Mining – KDD96, N.Y.: AAAI Press, Aug. 2*, 1996, pp. 226–231.
6. Hinneburg A., Klein D. An efficient approach to clustering in large multimedia databases with noise, *Proc. 4th Int. Conf. in Knowledge Discovering and Data Mining – KDD98, N.Y.: AAAI Press, Aug. 27, 1998.* Hinneburg, 1998, pp. 58–65.
7. Ankerst M., Brening M., Kriegel H., Sander J. OPTICS: Ordering points to identify the clustering structure. *Proc. 1999 ACM SIGMOD Int. Conf. on Management of Data, Jun. 1, 1999*. Philadelphia, 1999, pp.49–60.
8. Rodriguez A., Laio A. Clustering by fast seach and find of density peaks, *Science,* 2014, № 34, pp. 1492–1496.
9. Höppner F., Klawonn F., Kruse R., Runkler T. Fuzzy Clustering Analysis: Methods for Classification, Data Analisys and Image Recognition. Chichester, John Wiley &Sons, 1999, 300 p.
10. Zhou J., Wang Q., Hung C.-C., Yi X. Credibilistic clustering: the model and algorithms, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2015, Vol. 23, No. 4, pp. 545–564. DOI: https://doi.org/10.1142/S0218488515500245
11. Zhou J., Wang, Q., Hung C. C. Credibilistic clustering algorithms via alternating cluster estimation, *Journal of Intelligent Manufacturing,* 2017, Vol. 28, pp. 727–738. DOI: https://doi.org/10.1007/s10845-014-1004-6.
12. Begum N., Ulanova L., Wamg J., Klogh E. Accelerating dynamic time warping clustering with a novel admissible pruning strategy, *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 10, 2015.* Sydney, NSW, Australia, pp. 49–58. DOI: https://doi.org/10.1145/2783258.27 83286.
13. Shafronenko A., Bodyanskiy Ye., Klymova I., Holovin O. Online credibilistic fuzzy clustering of data using membership functions of special type[Electronic resource], *Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), April 27–1 May 2020.* Zaporizhzhia, 2020. Access mode: http://ceur-ws.org/Vol-2608/paper56.pdf.

DOI 10.15588/1607-3274-2022-1-9

e-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2022. № 1
p-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2022. № 1

UDC 004.93

FASTER OPTIMIZATION-BASED META-LEARNING ADAPTATION PHASE

Khabarlak K. S. – Post-graduate student of the Department of System Analysis and Control, Dnipro University of Technology, Dnipro, Ukraine.

ABSTRACT

Context. Neural networks require a large amount of annotated data to learn. Meta-learning algorithms propose a way to decrease number of training samples to only a few. One of the most prominent optimization-based meta-learning algorithms is MAML. However, its adaptation to new tasks is quite slow. The object of study is the process of meta-learning and adaptation phase as defined by the MAML algorithm.

Objective. The goal of this work is creation of an approach, which should make it possible to: 1) increase the execution speed of MAML adaptation phase; 2) improve MAML accuracy in certain cases. The testing results will be shown on a publicly available few-shot learning dataset CIFAR-FS.

Method. In this work an improvement to MAML meta-learning algorithm is proposed. Meta-learning procedure is defined in terms of tasks. In case of image classification problem, each task is to try to learn to classify images of new classes given only a few training examples. MAML defines 2 stages for the learning procedure: 1) adaptation to the new task; 2) meta-weights update. The whole training procedure requires Hessian computation, which makes the method computationally expensive. After being trained, the network will typically be used for adaptation to new tasks and the subsequent prediction on them. Thus, improving adaptation time is an important problem, which we focus on in this work. We introduce Λ (lambda) pattern by which we restrict which weight we update in the network during the adaptation phase. This approach allows us to skip certain gradient computations. The pattern is selected given an allowed quality degradation threshold parameter. Among the pattern that fit the criteria, the fastest pattern is then selected. However, as it is discussed later, quality improvement is also possible is certain cases by a careful pattern selection.

Results. The MAML algorithm with Λ pattern adaptation has been implemented, trained and tested on the open CIFAR-FS dataset. This makes our results easily reproducible.

Conclusions. The experiments conducted have shown that via Λ adaptation pattern selection, it is possible to significantly improve the MAML method in the following areas: adaptation time has been decreased by a factor of 3 with minimal accuracy loss. Interestingly, accuracy for one-step adaptation has been substantially improved by using Λ patterns as well. Prospects for further research are to investigate a way of a more robust automatic pattern selection scheme.

KEYWORDS: few-shot learning, meta-learning, Model-Agnostic Meta-Learning, MAML, adaptation time, adaptation speed, optimization-based meta-learning.

ABBREVIATIONS

MAML is Model-Agnostic Meta-Learning, a method of optimization-based few-shot learning;

ResNet is a Residual Network, a particular architecture of Convolutional Neural Networks;

NLP is Natural Language Processing.

NOMENCLATURE

N is a number of images per class that are given for the network training.

K is a number of classes the network is trained to distinguish between.

X is a network input, in our case images.

$\Phi(\theta, X)$ is a neural network.

θ is a matrix of network weights.

B is a number of layers in the neural network.

$\rho(T)$ is a distribution of all tasks.

T_i is one of the tasks, consisting of Support Set S_i, Query Set Q_i.

P is a number of adaptation steps.

$\theta_i^{(j)}$ is a matrix of adapted weights after j iterations that correspond to i^{th} task.

α is an adaptation step size, $\alpha > 0$.

β is a learning rate, $\beta > 0$.

Λ is an adaptation template, which controls which neural network layers should be updated during the adaptation procedure to the current task T.

INTRODUCTION

The neural network accuracy for image classification has significantly improved thanks to deep convolutional neural networks. However, a very large number of images is required for such networks to train successfully. For instance, all of the ResNet [1] neural network configurations from ResNet-18 to ResNet-152 (18 and 152 layers deep correspondingly) are trained on the ImageNet dataset [2], which contains 1.281.167 images and 1.000 classes (about 1.200 samples per class). Obviously, for many of the practically significant tasks it is impossible to collect and label a dataset that large. Thus, learning deep convolutional networks from scratch might yield poor results. Because-of that, on the smaller datasets typically an approach called transfer learning is used instead. That is, an ImageNet pretrained network of a particular architecture is taken and then further finetuned on the target (smaller) dataset [1; 3; 4]. However, training on few examples per class is still a challenge. This contrasts to how we, humans, learn, when even a single example given to a child might be enough. Also, it is hard to estimate the

DOI 10.15588/1607-3274-2022-1-10

quality of a certain ImageNet pretrained network on the target dataset. Hence, we get a model selection problem: if the model A is better than the model B on ImageNet, will it be better on our small dataset? A promising approach to resolving both of these problems is to use meta-learning or its benchmark known as few-shot learning. Meta-learning trains the network on a set of different tasks, which are randomly sampled from the whole space of tasks. By learning the network in such a way, it is assumed that the network will learn features that are relevant to all of the tasks and not only to the single one, i.e., will learn more general features.

In this work we focus on one of the most prominent optimization-based meta-learning methods, called MAML [5]. This method has become a keystone, and as it will be shown in the literature overview section, many of the newer method base on its ideas. Training of the MAML method is split into the so-called adaptation and meta-gradient update phases.

The subject of study is the class of optimization-based meta-learning algorithms.

It has been shown that adaptation phase of the MAML is quite slow to perform [6], and in general, high neural network execution speed is a major problem for applications [7]. In this work we introduce gradient update patterns, i.e., a selective update of the neural network weights during the adaptation phase.

The purpose of this work is to show that by carefully selecting the newly-proposed gradient update pattern, it is possible to: 1) increase the execution speed of MAML adaptation phase; 2) significantly improve MAML performance in case, when only 1 adaptation phase is used. The testing results will be shown on a publicly-available few-shot learning dataset CIFAR-FS [8].

1 PROBLEM STATEMENT

The goal behind meta-learning is to train a neural network $\Phi(\theta)$, that is capable of adapting to the new previously unknown tasks given a small number of examples. Meta-learning is also said to be learning to learn problem. The training procedure is defined using a concept of tasks, that are sampled from the whole task space $\rho(T)$ of the problem domain. The task is a tuple $T = \{S, Q\}$, consisting of the so-called Support Set $S = \{X_S, y_S\}$ and Query Set $Q = \{X_Q, y_Q\}$ [5; 9–11]. In literature, the Query Set is also sometimes referred to as Target Set. Support Set $\{X_S, y_S\}$ is used to adapt (or train) the network to the new task. The set S is small. X_S are the network inputs, y_S – the expected predictions. The number of examples per class is denoted as K and written as K-shot. K is typically in range from 1 to 20, although no hard upper-bound is defined. X_Q, y_Q are the query inputs and expected outputs correspondingly. Number of classes N the network should distinguish between is denoted as N-way.

We have given the general training procedure, next we define it in more detail for image classification optimization-based meta-learning, which this paper is focused on. Optimization meta learning is defined in 2 steps: 1) adaptation step, which computes adaptation weights in a form of function $\theta'(\theta)$, that minimize task-specific error $L(y_s, \Phi(\theta', X_S))$; 2) meta-gradient update, which updates meta-weights θ. The idea behind such training procedure is that by finding good weights θ, it will be possible to adapt to new previously unseen tasks with few training examples in the adaptation procedure. For classification, the loss function used is typically cross-entropy (1):

$$L(y, \Phi(\theta, X)) = -\sum_i y_i \log \Phi(\theta, X_i). \quad (1)$$

We define the algorithm-specific part in the Materials and Methods section. In this work we set a goal of improving adaptation step execution time and accuracy.

2 REVIEW OF THE LITERATURE

The meta-learning approaches are mainly divided into 3 broad categories [12]: metric-based, model-based and optimization-based. Representatives of each group differ in the neural network design and training procedure. In this work we focus on classification methods, yet applications exist in literally every field of machine learning [5; 13–15], such as NLP, Reinforcement Learning, Face Verification, etc.

Next, we describe each category of meta-learning methods. 1) In metric-based methods the goal is to define a neural network architecture that produces an embedding into a metric space and a similarity measure (metric), so that the distance between embeddings of the same class is smaller than that of different classes. Examples of such methods include Siamese Networks [16], Matching Networks [17], Prototypical Networks [9]. 2) In model-based methods the network architecture is designed, so that the model has explicit memory cells, which help the network to adapt quickly, for instance, Memory-Augmented Neural Networks [18]. 3) In optimization-based learning the network architecture is not changed, which means that conventional architectures for image classification can be used. One of the quintessential methods in this category is MAML [5], which defines the training procedure as a 2nd-order optimization problem. The method applicability has been shown in regression, classification and reinforcement learning. Two popular datasets were considered for image classification: Omniglot [19] and miniImageNet [10; 17], where MAML has beaten with a margin many of the previous methods. After MAML has been introduced, a lot of works have proposed its modifications. Reptile [20] has simplified MAML training scheme, MAML++ [11] has given practical recommendation on improving MAML training stability. In has been noted that while MAML++ has introduced more parameters to the network, total training time has decreased thanks to the performance optimizations proposed. Authors of Meta-SGD [21] note that by learning not only network weights, but also separate update coefficient for each of the weights, it is possible to achieve higher accuracies. However, the network training time and memory con-

DOI 10.15588/1607-3274-2022-1-10

sumption has significantly increased as twice the number of the parameters should be optimized.

In contrast to previous works, in this paper we focus on improving the network adaptation and not training time. We assume that after the initial training, the network can be adapted to multiple tasks in an online format. Thus, minimizing adaptation time is an important problem. The results obtained in the paper will be applicable to many of the optimization-based algorithms, including but not limited to the ones mentioned above.

3 MATERIALS AND METHODS

In this work we propose a modification to the MAML algorithm. As we have described in the problem statement section above, this class of algorithms is defined in terms of adaptation and meta-gradient update phases.

The algorithm starts by randomly sampling a training task $T_i \sim \rho(T)$. To sample a task T_i means to 1) randomly select N classes from all classes that are available in the dataset split (training, validation or test, based on which accuracy we want to compute); 2) randomly select K images per each of N classes for the Support Set and K_Q images per each class N for the Query Set. The first phase of the algorithm is adaptation, where MAML minimizes loss function (1) on the Support Set by performing several stochastic gradient descent steps. To do that the algorithm iteratively builds model weights $\theta_i^{(j)}(\theta)$ via formula (2), note that $\theta_i^{(0)} \equiv \theta$:

$$\theta_i^{(j)} = \theta_i^{(j-1)} - \alpha \nabla_{\theta_i} L\left(y_{S_i}, \Phi\left(\theta_i^{(j-1)}, X_{S_i}\right)\right) \quad (2)$$

Having iteratively built the task specific weights $\theta_i^{(j)}$, the algorithm updates the meta-weights θ using formula (3):

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{Q_i \in T_i} L\left(y_{Q_i}, \Phi\left(\theta_i^{(P)}, X_{Q_i}\right)\right) \quad (3)$$

In essence, in (3) the algorithm updates the meta-weights θ by averaging computed loss function (1) on the Query Set, for the neural networks Φ with weights $\theta_i^{(P)}$ on several tasks T_i, i.e., in this step the algorithm backpropagates through the losses of all the task-specific adaptations. Throughout the paper we use 4 tasks for the meta-update step. Note, that in (2) task-specific weights $\theta_i^{(j)}$ are computed on the Support Set, and in (3) Query Set is used for the loss computation. Also, in contrast to the conventional neural network training procedure the loss function is computed twice: first, to compute the adaptation weights $\theta_i^{(P)}$ in (2); second, to compute the resulting adaption loss in (3). Also, in (2) the gradient is taken by task-specific weights $\theta_i^{(j-1)}$ from previous step, and in (3) the gradient is taken by meta-weights θ. Thus, as can be seen from formulas (2), (3), the method requires Hessian computation during the meta-gradient update, hence, this is a second-order optimization method. The whole training procedure can be seen in algorithm 1. A more detailed information can be found in the original paper [5].

Algorithm 1. MAML adaptation procedure

```
Randomly sample task T_i from task space ρ(T)
For each task T_i={S_i, Q_i}, where S_i = {X_S, y_S}, Q_i = {X_Q, y_Q}
    For iteration j = {1, …, P}
        Adapt the network via formula (2) using S_i
    End for
End for
Update meta-weight θ via (3) using Q_i and the task specific weights θ_i^(P)
```

Next, we define our modified adaptation procedure. Given a convolutional neural network that has B layers, we define an adaptation pattern (4), where Λ_j is an indicative function as defined in (5), which indicates layers of the network that should be updated during backpropagation.

$$\Lambda = \{\Lambda_1, \Lambda_2, ..., \Lambda_B\}, \quad (4)$$

$$\forall l : \Lambda_l = \begin{cases} 1, \textit{ if layer } l \textit{ is updated}, \\ 0, \textit{ otherwise}. \end{cases} \quad (5)$$

We say that pattern is full if $\forall l : \Lambda_l = 1$. In this case our adaptation phase will be equivalent to the one proposed in MAML. We consider all possible patterns Λ, except $\forall l : \Lambda_l = 0$, when no weights can be updated, thus, no adaptation is possible. We assume that updating only certain weights might be useful, because the neural networks tend to learn features that differ in complexity, the closer the layer is to the input the simple the features are [22]. Also, authors of Meta-SGD [21] have shown that by learning weight-specific learning rates the resulting quality was superior to the original MAML algorithm. However, Meta-SGD approach was much slower to train as both weights and learning rates have to be learned during the training procedure. Training time in our approach is intact. In contrast to previous works, we propose to update only certain weights, thus, essentially freezing some layers. This allows us to decrease gradient computations required during the adaptation phase as is shown on Fig. 1 for a convolutional network that contains 4 convolutional and a single fully-connected (linear) layer.

In Fig. 1 the backpropagation pass goes in the direction opposite to arrows (forward pass). The architecture is taken as an example and can be arbitrary in practice. For the example pattern $\Lambda = \{0,1,0,1,1\}$, we can see that for the Convolutional Block 4 and the Linear layers both the gradient is computed and the weights are updated. For Convolutional Block 3 gradients are computed, but weights are not updated as Convolutional Block 2 requires weight update. However, for Convolutional Block 1 no gradients computation or weight update are performed.

DOI 10.15588/1607-3274-2022-1-10

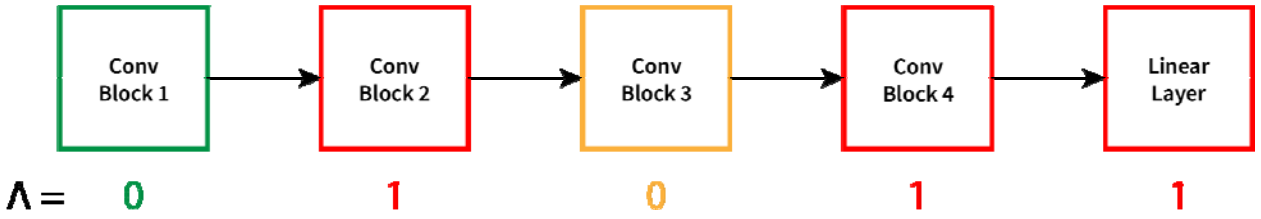


Figure 1 – Λ pattern backpropagation scheme. Backpropagation is performed in order reverse to the arrows. In red – gradients are computed, networks weights are updated; yellow – gradients are computed, no network weight update; green – both gradient computation and network weight update are skipped

Given the above-described Λ pattern description, the updated adaptation formula will look as follows (6):

$$\theta_i^{(j)} = \theta_i^{(j-1)} - \Lambda\alpha\nabla_{\theta_i} L\left(y_{S_i}, \Phi\left(\theta_i^{(j-1)}, X_{S_i}\right)\right) \qquad (6)$$

4 EXPERIMENTS

To conduct the experiments, we have reimplemented the MAML algorithm. The following paragraph describes the details.

The authors of MAML have defined convolutional neural network architecture and have used it for miniImageNet experiments. This network is commonly referred to as "CNN4" in the later meta-learning literature. It has 4 convolutional blocks, followed by a linear layer. Each of the blocks has a convolutional layer with kernel size of 3 and padding of 1, followed by the Batch Normalization [23], ReLU activation and Max Pooling with kernel size of 2. Number of filters in the convolutional layers is a configurable parameter, the authors have used 32, which we follow. Number of outputs in the linear layer is defined by K for K-way classification problem. Training is performed via Adam [24] gradient descent method as meta-optimizer with learning rate of $\beta = 10^{-3}$ and $\alpha = 0.01$ as the adaptation step size. Each model has been trained for 600 epochs. While the authors used meta-batch size of 2 for 5-shot and 4 for 2-shot experiment to reduce training memory consumption, we stick to 4 as it leads to slightly better performance on CIFAR-FS [8] dataset during our experiments. Also, the dataset memory footprint is small, so we don't have to reduce memory consumption by using a smaller batch-size. Each epoch has 100 randomly sampled tasks. For the gradient update $N \cdot K$ samples are taken for N-shot K-way classification problem for training and 15 samples per class for evaluation, thus following [10].

In addition, we have modified the network adaptation procedure, so that it updates only weights defined by pattern Λ as defined in (4)–(6).

For the experiments we have used the novel CIFAR-FS [8] dataset. It has been constructed from a well-known classification dataset, called CIFAR-100 [25]. It has images of different kinds of mammals, reptiles, flowers, man-made things, etc. The images are in color and have a size 32x32. Originally, this dataset was not supposed to be used in a few-shot learning setting. In [8] it has been suggested to split 100 classes into train, validation and test sets. If it has been the non-few-shot neural network training, we would expect all of the 100 classes to be represented in each of the sets, only the images themselves would have been split. However, in few-shot learning case different disjoint classes are taken. Thus, 64 training, 16 validation and 20 test set classes have been selected. The exact classes that go into each split are important for testing the resulting accuracy and are defined in [8]. By using different classes for training and testing, the adaptation to the new classes can be better estimated. After such training the model is expected to quickly adapt to the new, unseen classes. We have taken the CIFAR-FS dataset for our experiments as it hasn't been analyzed by the MAML authors and is also faster to compute than miniImageNet.

All of the training procedures and time measurements were done on our own MAML implementation and tested on NVIDIA GTX 1050Ti GPU.

5 RESULTS

Given the network configuration as described in the experiments section, we have implemented the MAML algorithm. CIFAR-FS accuracy and adaptation timings are presented in Table 1.

Table 1 – Accuracies and adaptation timings on CIFAR-FS dataset

	1-shot 2-way	5-shot 2-way	1-shot 5-way	5-shot 5-way
Accuracy	77.2%	87.6%	51.7%	70.3%
Time	38.43 ms	40.70 ms	41.67 ms	45.35 ms

In (6) we have proposed a modified adaptation scheme, where only a part of weights is updated during the adaptation procedure. To begin with, we consider only trivial patterns Λ, where only one network layer is updated during the adaptation procedure. We show the accuracy on the test set in Fig. 2, where in a and b we conduct the experiment for 1-shot 5-way and 5-shot 5-way configurations correspondingly. To see the impact of the number of adaptation steps, we also show the accuracies for $P = 10$ (default) and 1, 3, 5 adaptation steps. As it can be seen, the model accuracy differs significantly between the configurations. For 1-shot 5-way, learning one of the three first convolutional layers only has no effect, the accuracy remains on the level of random guessing (20%). However, training either convolutional layer 4 or the last linear layer improves the model accuracy. Note, that the number of parameters in layers differs. In Table 2, we

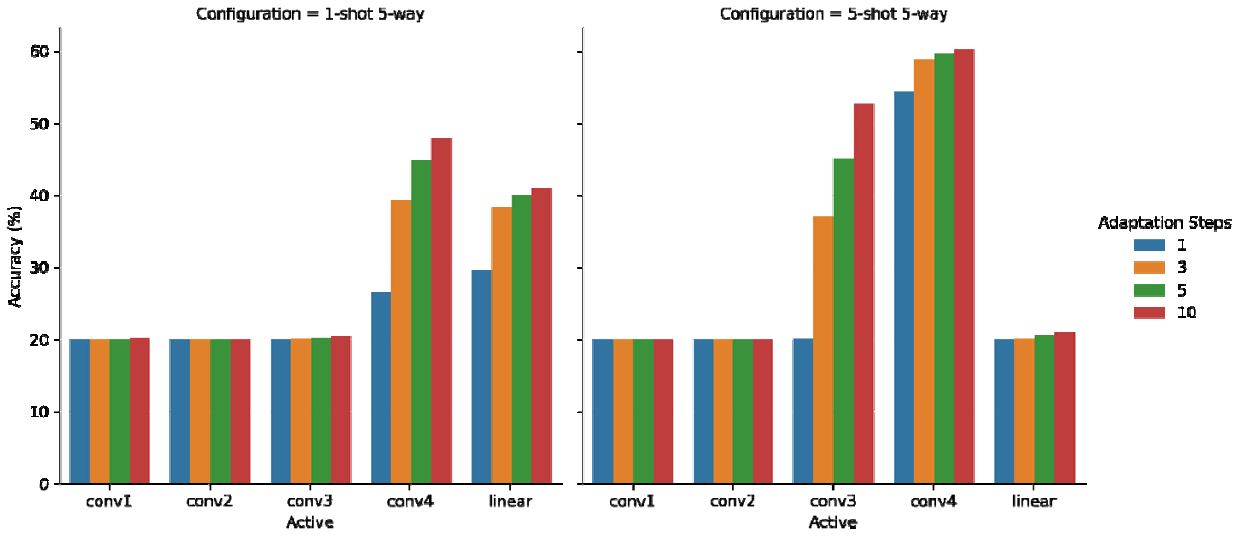
DOI 10.15588/1607-3274-2022-1-10

show the number of parameters for each layer. Note that final layer has different number of parameters depending on *N* output classes. It can be seen that the first convolutional layer and the final linear (fully-connected) layers have fewer parameter than inner convolutional blocks. This can explain the fact that learning only the linear layer has worse performance. For 5-shot 5-way we see that only convolutional layers 3 and 4 have a positive impact on the performance if adapted alone. Interestingly, number of adaptation steps has a significant impact on the performance with only convolutional layer #3 enabled. As we will see later, such an impact is higher, than when the full network is updated during the adaptation.

In Fig. 3, a and b we depict a similar experiment for 1-shot 2-way and 5-shot 2-way configurations correspondingly. Note, that random guessing baseline for these configurations is now at 50%, so the lower bound for accuracy is now higher than in Fig. 2. Here we see an opposite trend, where updating the first layers also has a positive impact on the resulting accuracy. Contrasting to previous experiment, updating the Convolutional Block 4 only doesn't provide the best results in either case.

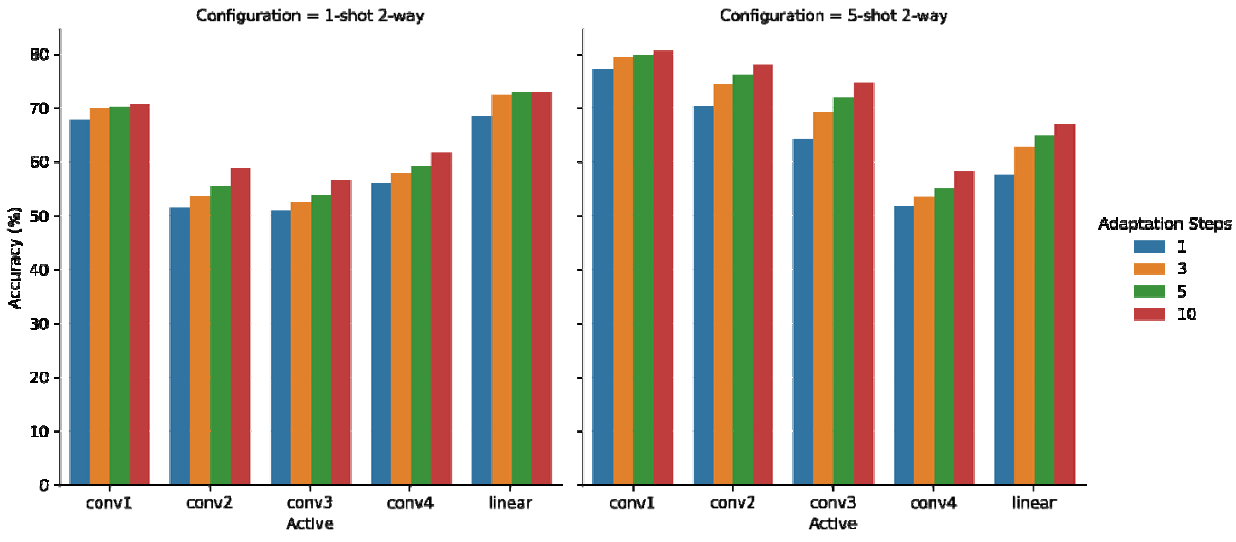
Table 2 – Number of parameters for each layer

Layer Name	Number of Parameters
Conv Block 1	960
Conv Block 2	9.312
Conv Block 3	9.312
Conv Block 4	9.312
Linear	1.602 (2-way) 4.005 (5-way)
Total	30.498 (2-way) 32.901 (5-way)



a b

Figure 2 – Adaptation accuracy for trivial Λ patterns, i.e., only a single layer is updated during adaptation: a is 1-shot 5-way, b is 5-shot 5-way



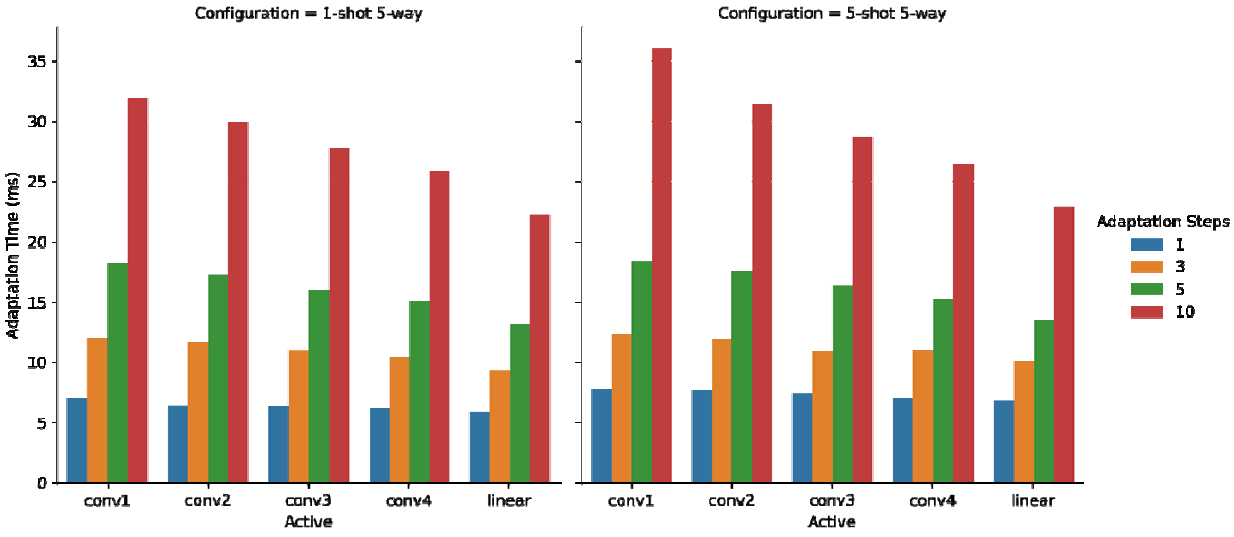
a b

Figure 3 – Adaptation accuracy for trivial Λ patterns, i.e., only a single layer is updated during adaptation: a is 1-shot 2-way, b is 5-shot 2-way

DOI 10.15588/1607-3274-2022-1-10

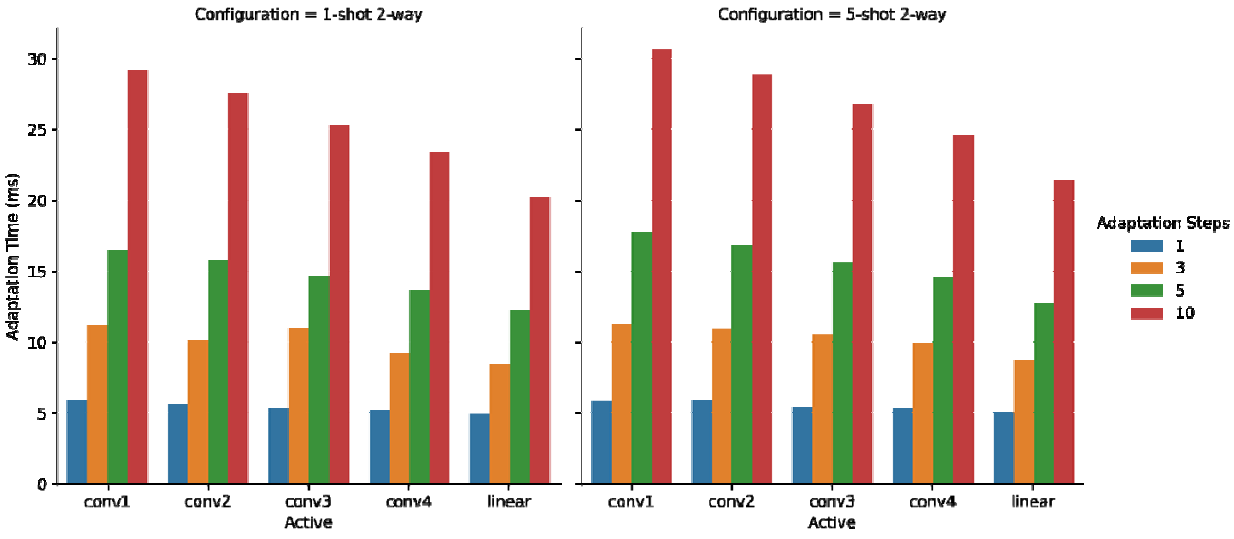
As one of the goals in our work is to improve the model adaptation speed, we have timed experiments for trivial patterns Λ. On Fig. 4 and 5 we show the model adaptation time corresponding to all of the four configurations depicted on Fig. 2 and 3. As we can see, in both cases we have a similar trend where the closer the layer we update to the end of the network, the smaller the adaptation time is. This follows our previous idea that by skipping some gradient computations (as have been shown on Fig. 1), adaptation time can be reduced.

As can be seen from Fig. 4 and 5, number of adaptation steps has a significant impact on the adaptation speed. On Fig. 6 we show the model accuracy for each of the four scenarios and on Fig. 7 we depict the corresponding timings, both shown with respect to the number of the adaptation steps. As before, the experiments have been conducted for $P = 1, 3, 5$ and 10 adaptation steps. The results between those reference points have been linearly interpolated. The presented accuracies and timings are the average taken for all 31 possible patterns Λ. Note, that throughout the article we exclude pattern $\forall l : \Lambda_l = 0$, as no weights can be changed for such pattern, therefore no adaptation is possible. As can be seen, while the adaptation time grows linearly with the number of adaptation steps, the accuracy growth plateaus at around 5 adaptation steps. Actually, for the full pattern Λ increasing number of adaptation steps from 5 to 10 has less than 0.3% improvement in accuracy. In typical practical scenarios such an improvement is insignificant. Thus, we suggest that performing 10 adaptation steps is redundant.



a b

Figure 4 – Adaptation time for trivial Λ patterns:
a is 1-shot 5-way, b is 5-shot 5-way



a b

Figure 5 – Adaptation time for trivial Λ patterns:
a is 1-shot 2-way, b is 5-shot 2-way

DOI 10.15588/1607-3274-2022-1-10

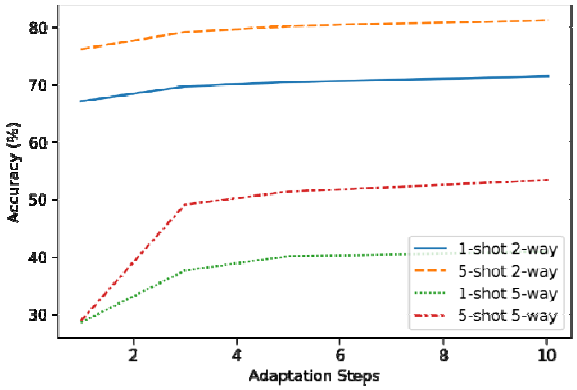


Figure 6 – Accuracy averaged for all patterns Λ for different *N*-shot *K*-way problems with respect to the number of adaptation steps *P*

Next, we try to search for such a pattern Λ and number of adaptation steps, so that the resulting accuracy drops no more than 0.07 times the full pattern accuracy. We see such a quality degradation threshold reasonable for practical applications. It should be noted that the approach we propose can be applied with an arbitrary quality degradation threshold. We show such patterns in table 3. Based on this table, we suggest using the $\Lambda^* = \{1,0,1,1,1\}$, which offers factor of 3.0 speed improvement with an insignificant quality loss. It can be seen that pattern $\Lambda = \{0,1,1,1,1\}$ also suits the specified criteria and also has a slightly higher (factor of 3.1) performance improvement, however, it has a significantly lower performance for both of the 2-way configurations, degrading on 2.5% and 3.2% relative to the best selected pattern Λ^*. We consider such a degradation not worth the speed up. The fact that enabling first CNN layer is significant for the 2-way learning accuracy, closely follows the presented above description of Fig. 3. Also, not to be mistaken, in Fig. 2–5, we had only one layer updated during the adaptation phase (thus $\Sigma_l \Lambda_l = 1$), however, the best selected pattern Λ^* has all expect one layer updated.

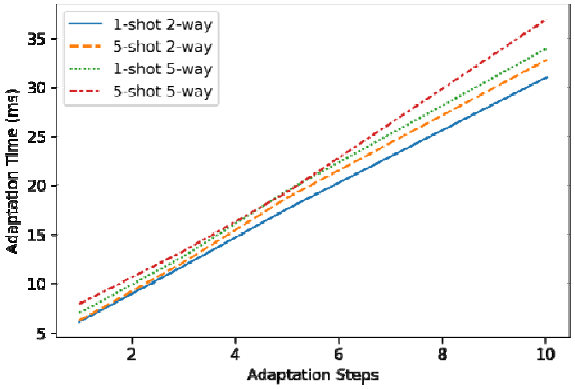


Figure 7 – Adaptation time averaged for all patterns Λ for different *N*-shot *K*-way problems with respect to the number of adaptation steps *P*

Finally, we pose a question, whether updating only part of weights in the neural network can improve the method performance. We have discovered, that in extreme case of learning with a single adaptation step ($P = 1$), we have significant improvement in 5-way adaptation performance by updating with a partial pattern Λ. The performance for the full pattern, as well as a partial, is shown in table 4.

Table 3 – Adaptation speedup depending on pattern Λ and the number of adaptation steps. Patterns with loss degradation of less than 7% (relative to full pattern Λ and 10 adaptation steps) are shown.

Adaptation Steps	Pattern Λ	1-shot 2-way (%)	5-shot 2-way (%)	1-shot 5-way (%)	5-shot 5-way (%)	Mean Adaptation Time (ms)	Relative Speedup (times)
3	0,1,1,1,1	74.7	83.2	49.3	69.7	13.3	3.1
3	1,0,1,1,1	76.6	85.9	49.3	69.8	13.9	3.0
3	1,1,1,1,1	76.6	87.2	49.3	70.0	15.0	2.8
5	0,1,1,1,1	75.2	83.9	51.5	69.9	20.0	2.1
5	1,0,1,1,1	76.9	86.2	51.4	70.1	21.1	2.0
5	1,1,1,1,1	77.0	87.4	51.6	70.2	22.6	1.8
10	0,1,1,1,1	75.4	84.6	51.7	70.1	36.1	1.2
10	1,0,1,1,1	77.1	86.6	51.7	70.1	38.6	1.1
10	1,1,1,1,1	77.2	87.6	51.7	70.3	41.5	1.0

Table 4 – Accuracy improvement for $P = 1$ gradient step adaptation with pattern selection

Accuracy	1-shot 2-way	5-shot 2-way	1-shot 5-way	5-shot 5-way
$\Lambda = \{1,1,1,1,1\}$	74.3%	86.0%	36.8%	20.4%
$\Lambda = \{1,1,0,1,1\}$	74.3%	83.1%	36.9%	53.1%

We have also performed a search of all cases, when our approach gives better results than the original with $P = 1$. The results are shown in Table 5.

DOI 10.15588/1607-3274-2022-1-10

Table 5 – Accuracy improvement for $P = 1$ gradient step adaptation with pattern selection if pattern is selected per configuration

	1-shot 2-way	5-shot 2-way	1-shot 5-way	5-shot 5-way
Accuracy on $\Lambda = \{1,1,1,1,1\}$	74.3%	86.0%	36.8%	20.4%
Accuracy on selected Λ	74.5%	86.2%	36.9%	54.8%
Selected Pattern Λ	1,1,1,0,1	1,1,1,0,1	1,1,0,1,1	1,1,0,1,0

6 DISCUSSION

In [22] it has been shown that each trained neural network's convolutional layer has a different meaning. The first layer tends to learn simple features, like edges, lines or color gradients. The second layer increases its complexity and understands simple shapes, e.g., circles, corners or stripes, while the last layers learn high-level features, such as eyes, faces, text-like objects, etc. The exact features learned, obviously, depend on the training dataset, however, such logic is retained. In the few-shot learning classification scenario the tasks differ by the types of objects that the model has to classify (e.g., horse, vehicle, frog, etc.). As we have described in the experiments section, train and test sets have different disjoint classes presented. Thus, it might be reasonable to expect that only the last layers of the network should be changed to adapt to the new tasks and classes. This is exactly what we see in the case of 5-way classification as is shown on Fig. 2. However, such a statement contradicts to the experiment results from Fig. 3. By examining the original CIFAR-100 dataset, we can see that image labels (classes) form larger coarse groups. For instance, coarse class (or superclass) "aquatic mammals" contains "beaver", "dolphin", "otter", "seal", "whale". Other examples of superclasses include "fish", "large carnivores", "household electrical devices", etc. The training itself is performed on finer classes. From the examples we have picked, it becomes obvious that instances of different classes have a significant variation in color. Images of aquatic mammals and fish typically contain blue and gray colors, while large carnivores might have more yellow and green. Thus, in case of 2-way classification it is more probable that both classes will be picked from a single or several similar superclasses than in case of 5-way classification. Consequently, we suggest that updating the first layer of a neural network in a 2-way few-shot learning scenario adjusts the feature distribution to the one expected by the following neural network layers. We see this as an analogy of how a human eye works: it adjusts the amount of light coming to the retina by expanding or contracting the pupil, so that it becomes easier to see the details.

From Table 3 we see that keeping the inner layers stale is the most fruitful way to improve the performance, with little to no quality loss. A substantial increase in adaptation speed has been achieved with a target quality loss set to 7% relative to the original pattern $\Lambda = \{1,1,1,1,1\}$ and $P = 10$ adaptation steps. The actual quality loss turns out to be even smaller as we have skipped slightly faster, but worse pattern $\Lambda = \{0,1,1,1,1\}$. Thereby, with the best $\Lambda^* = \{1,0,1,1,1\}$ and $P = 3$ adaptation steps, we achieve a factor of 3.0 speed improvement. Our quality losses are the following: 1-shot 2-way is 0.78%, 5-shot 2-way is 1.97% 1-shot 5-way is 4.86% and 5-shot 5-way is 0.71%. Even smaller quality losses can be achieved by consulting table 3. Note, that these are relative quality losses. If the losses are computed in absolute terms, they become even more negligible. Thus, we state that have achieved a significant adaptation time reduction with small-enough quality loss.

We also discuss a way to improve algorithm quality by selecting a pattern Λ. In an extreme case of single adaptation step, avoiding to update the inner layer has helped to improve the overall model quality as is shown in table 4. We have also been able to find such a pattern for each of the few-shot learning configurations such that it improves the model performance for $P = 1$ adaptation step in table 5. It is curious that no such behavior is observed in cases when $P > 1$. To the best of our knowledge such behavior has not been previously observed and should be further investigated.

CONCLUSIONS

MAML is an optimization-based few-shot learning method that is able to learn an arbitrary neural network by using only a few samples per class. Many algorithms follow the learning scheme proposed in MAML. In this work we solve the problems of 1) long adaptation time, and 2) poor performance in cases when a single adaptation step is used.

The scientifical novelty of obtained results is that the method of reducing number of gradient computations during MAML adaptation phase has been introduced via the newly proposed Λ patterns. By selecting an appropriate adaptation pattern, we have significantly improved the method in the following areas: 1) long MAML adaptation time has been decreased by the factor 3 with minimal accuracy loss; 2) accuracy for cases when only a single adaptation step is used has been substantially improved.

The practical significance of obtained results is that an improvement of adaptation time of the widespread MAML algorithm will enable applicability of the algorithm on less powerful devices and will in general decrease the time needed for the algorithm to adapt to new tasks.

Prospects for further research are to investigate a way of a more robust automatic pattern selection scheme for an arbitrary training dataset and network configuration.

ACKNOWLEDGEMENTS

The author expresses gratitude to Larysa Koriashkina, PhD, Associate Professor of the Department of System Analysis and Control, Dnipro University of Technology for support and a fruitful paper discussion. The work is supported by the state budget scientific research project of Dnipro University of Technology "Development of new

DOI 10.15588/1607-3274-2022-1-10

mobile information technologies for person identification and object classification in the surrounding environment" (state registration number 0121U109787).

REFERENCES

1. He K., Zhang X. , Ren S. et al. Deep Residual Learning for Image Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR 2016. Las Vegas, NV, USA, June 27–30, 2016. IEEE Computer Society, 2016. pp. 770–778. DOI: 10.1109/CVPR.2016.90.
2. Deng J., Dong W., Socher R. et al. ImageNet: A large-scale hierarchical image database, *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20–25 June 2009*. Miami, Florida, USA, IEEE Computer Society, 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
3. Huang G., Liu Z., Maaten L. et al. Densely Connected Convolutional Networks, 2*017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. Honolulu, HI, USA, July 21–26, 2017, IEEE Computer Society, 2017, pp. 2261–2269. DOI: 10.1109/CVPR.2017.243.
4. Zagoruyko S., Komodakis N. Wide Residual Networks, *Proceedings of the British Machine Vision Conference (BMVC)*. BMVA Press, 2016, pp. 87.1–87.12. DOI: 10.5244/C.30.87.
5. Finn C., Abbeel P., Levine S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6–11 August 2017, Proceedings of Machine Learning Research*. PMLR, 2017, Vol. 70, pp. 1126–1135.
6. Rajeswaran A., Finn C., Kakade S. et al. Meta-Learning with Implicit Gradients, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019*. Vancouver, BC, Canada, 2019, pp. 113–124.
7. Khabarlak K., Koriashkina L. Fast Facial Landmark Detection and Applications: A Survey [Electronic resource], *arXiv:2101.10808 [cs]*, 2021. Access mode: https://arxiv.org/abs/2101.10808
8. [Bertinetto L., Henriques J., Torr P. et al. Meta-learning with differentiable closed-form solvers [Electronic resource], *7th International Conference on Learning Representations, ICLR 2019*. New Orleans, LA, USA, May 6–9, 2019. Access mode: https://openreview.net/forum?id=HyxnZh0ct7.
9. Snell J., Swersky K., Zemel R.S. // Prototypical Networks for Few-shot Learning, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, December 4–9, 2017. Long Beach, CA, USA, 2017, pp. 4077–4087.
10. Ravi S., Larochelle H. Optimization as a Model for Few-Shot Learning [Electronic resource], *5th International Conference on Learning Representations, ICLR 2017*. Toulon, France, April 24–26, 2017, Conference Track Proceedings. Access mode: https://openreview.net/forum?id=rJY0-Kcll.
11. Antoniou A., Edwards H., Storkey A. J. How to train your MAML [Electronic resource], *7th International Conference on Learning Representations, ICLR 2019*. New Orleans, LA, USA, May 6–9, 2019. Access mode: https://openreview.net/forum?id=HJGven05Y7.
12. Weng L. Meta-Learning: Learning to Learn Fast [Electronic resource]. Access mode: https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html.
13. Yin W. Meta-learning for Few-shot Natural Language Processing: A Survey [Electronic resource], *CoRR*, 2020, Vol. abs/2007.09604. Access mode: https://arxiv.org/abs/2007.09604.
14. Wang Y., Yao Q. , Kwok J. et al. Generalizing from a Few Examples: A Survey on Few-shot Learning, *ACM Comput. Surv*, 2020, Vol. 53, No. 3, pp. 63:1–63:34. DOI: 10.1145/3386252.
15. Guo Y., Zhang L. One-shot Face Recognition by Promoting Underrepresented Classes [Electronic resource], *CoRR*, 2017, Vol. abs/1707.05574. Access mode: http://arxiv.org/abs/1707.05574.
16. Koch G., Zemel R., Salakhutdinov R. Siamese neural networks for one-shot image recognition / G. Koch, // *ICML deep learning workshop*. Lille, 2015, Vol. 2.
17. Vinyals O., Blundell C., Lillicrap T. et al. Matching Networks for One Shot Learning, *Advances in Neural Information Processing Systems 29, Annual Conference on Neural Information Processing Systems 2016, December 5–10, 2016*. Barcelona, Spain, 2016, pp. 3630–3638.
18. Santoro A., Bartunov S., Botvinick M. et al. Meta-Learning with Memory-Augmented Neural Networks, *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19–24, 2016, JMLR Workshop and Conference Proceedings. JMLR.org*, 2016, Vol. 48, pp. 1842–1850.
19. Lake B. M., Salakhutdinov R., Tenenbaum J. B. Human-level concept learning through probabilistic program induction, *Science*, 2015, Vol. 350, No. 6266, pp. 1332–1338. DOI: 10.1126/science.aab3050.
20. Nichol A., Achiam J., Schulman J. On First-Order Meta-Learning Algorithms [Electronic resource], *CoRR*, 2018, Vol. abs/1803.02999. Access mode: http://arxiv.org/abs/1803.02999.
21. Li Z., Zhou F., Chen F. et al. Meta-SGD: Learning to Learn Quickly for Few Shot Learning [Electronic resource], *CoRR*. 2017, Vol. abs/1707.09835. Access mode: http://arxiv.org/abs/1707.09835.
22. Zeiler M.D., Fergus R. Visualizing and Understanding Convolutional Networks, *Computer Vision, ECCV 2014, 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part I, Lecture Notes in Computer Science*. Springer, 2014, Vol. 8689, pp. 818–833. DOI: 10.1007/978-3-319-10590-1_53.
23. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6–11 July 2015: JMLR Workshop and Conference Proceedings. – JMLR.org*, 2015, Vol. 37, pp. 448–456.
24. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization, *3rd International Conference on Learning Representations, ICLR 2015*. San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings, 2015.
25. Krizhevsky A. Learning multiple layers of features from tiny images [Electronic resource], *University of Toronto*, 2009, Access mode: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf

Received 22.11.2021.
Accepted 14.01.2022.

DOI 10.15588/1607-3274-2022-1-10

УДК 004.93

ПРИСКОРЕННЯ ФАЗИ АДАПТАЦІЇ ОПТИМІЗАЦІЙНОГО МЕТА-НАВЧАННЯ

Хабарлак К. С. – аспірант кафедри Системного аналізу та управління Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.

АНОТАЦІЯ

Актуальність. Нейронні мережі потребують багато розмічених даних для навчання. Алгоритми мета-навчання пропонують спосіб навчатися лише за декількома прикладами. Один з найзначніших алгоритмів оптимізаційного мета-навчання – це MAML. Однак, його процедура адаптації до нових задач є досить повільною. Об'єктом дослідження є процес мета-навчання та фаза адаптації в тому вигляді, як її визначено в алгоритмі MAML.

Мета. Метою даної роботи є створення підходу, що дозволить: 1) зменшити час виконання адаптації алгоритму MAML; 2) покращити якість алгоритму в ряді випадків. Показати результати тестування на публічно доступному наборі даних для мета-навчання CIFAR-FS.

Метод. В даній роботі запропоновано покращення алгоритму мета-навчання MAML. Процедура мета-навчання визначається через так звані «задачі». В разі класифікації зображень кожна задача є спробою навчитися класифікувати зображення нових класів лише за декількома навчальними прикладами. В алгоритмі MAML визначено 2 кроки процедури навчання: 1) адаптація до нової задачі; 2) оновлення мета-параметрів мережі. Вся тренувальна процедура потребує обчислення гесіану, що робить метод обчислювально складним. Після навчання мережа, зазвичай, буде використовуватися для адаптації до нових задач та наступної класифікації на них. Таким чином, покращення часу адаптації мережі є важливою проблемою. Саме на цій проблемі ми фокусуємося в даній роботи. Нами запропоновано шаблон Λ (лямбда) за допомогою якого ми обмежуємо, які параметри мережі слід оновлювати під час кроку адаптації. Даний підхід дозволяє не обчислювати градієнти для обраних параметрів та таким чином зменшити кількість необхідних обчислень. Шаблон обирається в межах параметру дозволеного зменшення якості мережі. Серед шаблонів, що відповідають заданому критерію, обирається найшвидший. Однак, як буде показано далі, в деяких випадках також можливе підвищення якості за допомогою правильно обраного шаблону адаптації.

Результати. Було реалізовано, навчено та перевірено якість роботи алгоритму MAML із шаблоном адаптації Λ на відкритому наборі даних CIFAR-FS, що робить отримані результати легко відтворюваними.

Висновки. Проведені експерименти показують, що із вибором шаблону Λ можливе значне покращення методу MAML в наступних областях: час адаптації було зменшено в 3 рази за мінімальних втрат якості. Цікаво, що для однокрокової адаптації якість значно виросла за умови використання запропонованого шаблону. Перспективи подальших досліджень можуть полягати в розробці більш робастного методу автоматичного вибору шаблонів.

КЛЮЧОВІ СЛОВА: пристрілкове навчання, мета-навчання, Model-Agnostic Meta-Learning, MAML, час адаптації, швидкість адаптації, оптимізаційне мета-навчання.

УДК 004.93

УСКОРЕНИЕ ФАЗЫ АДАПТАЦИИ ОПТИМИЗАЦИОННОГО МЕТА-ОБУЧЕНИЯ

Хабарлак К. С. – аспирант кафедры Системного анализа и управления Национального технического университета «Днепровская политехника», Днепр, Украина.

АННОТАЦИЯ

Актуальность. Нейронные сети требуют большого количества размеченных данных для обучения. Алгоритмы мета-обучения предлагают способ обучаться лишь по нескольким примерам. Одним из наиболее выдающихся алгоритмов оптимизационного мета-обучения является MAML. Однако, его процедура адаптации к новым задачам достаточно медленная. Объектом исследования является процесс мета-обучения и фаза адаптации в виде, как она определена в алгоритме MAML.

Цель. Цель данной работы – создание подхода, которых позволит: 1) уменьшить время выполнения адаптации алгоритма MAML; 2) улучшить качество алгоритма в ряде случаев. Показать результаты тестирования на открытом наборе данных для мета-обучения CIFAR-FS.

Метод. В данной работе предложено улучшение алгоритма мета-обучения MAML. Процедура мета-обучения определяется через так называемые «задачи». В случае классификации изображений каждая задача является попыткой научиться классифицировать изображения новых классов по нескольким обучающим примерам. В алгоритме MAML определено 2 шага в процедуре обучения: 1) адаптация к новой задаче; 2) обновления мета-параметров сети. Вся процедура обучения требует вычисление гессиана, что делает метод вычислительно сложным. После обучения сеть, как правило, будет использоваться для адаптации к новым задач и последующей классификации на них. Таким образом, улучшение времени адаптации сети является важной проблемой. Именно на этой проблеме мы и фокусируемся в данной работе. Нами предложено шаблон Λ (лямбда), с помощью которого мы ограничиваем, какие параметры сети следует обновлять во время шага адаптации. Данный подход позволяет не вычислять градиенты для выбранных параметров и таким образом уменьшить количество необходимых вычислений. Шаблон выбирается в рамках значения параметра разрешенного падения качества сети. Среди шаблонов, которые соответствуют заданному критерию, выбирается наиболее быстрый. Однако, як будет показано дальше, в некоторых случаях также возможно повышение качества с помощью правильно выбранного шаблона адаптации.

Результаты. Было реализовано, обучено и проверено качество работы алгоритма MAML с шаблоном адаптации Λ на открытом наборе данных CIFAR-FS, что делает полученные результаты легко воспроизводимыми.

Выводы. Проведенные эксперименты показывают, что с выбором шаблона Λ возможно значительное улучшение метода MAML в следующих областях: время адаптации было уменьшено в 3 раза при минимальных потерях в качестве. Интересно и то, что для одношаговой адаптации качество значительно выросло при условии использования выбранного шабло-

DOI 10.15588/1607-3274-2022-1-10

на. Перспективы дальнейших исследований могут заключаться в разработке более робастного метода автоматического выбора шаблонов.

КЛЮЧЕВЫЕ СЛОВА: пристрелочное обучение, мета-обучение, Model-Agnostic Meta-Learning, MAML, время адаптации, скорость адаптации, оптимизационное мета-обучение.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Deep Residual Learning for Image Recognition / [K. He, X. Zhang, S. Ren et al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27–30, 2016. – IEEE Computer Society, 2016. – P. 770–778. DOI: 10.1109/CVPR.2016.90.
2. ImageNet: A large-scale hierarchical image database / [J. Deng, W. Dong, R. Socher et al.] // 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20–25 June 2009, Miami, Florida, USA. – IEEE Computer Society, 2009. – P. 248–255. DOI: 10.1109/CVPR.2009.5206848.
3. Densely Connected Convolutional Networks / [G. Huang, Z. Liu, L. Maaten et al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21–26, 2017. – IEEE Computer Society, 2017. – P. 2261–2269. DOI: 10.1109/CVPR.2017.243.
4. Zagoruyko S. Wide Residual Networks / S. Zagoruyko, N. Komodakis // Proceedings of the British Machine Vision Conference (BMVC). – BMVA Press, 2016. – P. 87.1–87.12. DOI: 10.5244/C.30.87.
5. Finn C. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks / C. Finn, P. Abbeel, S. Levine // Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6–11 August 2017: Proceedings of Machine Learning Research. – PMLR, 2017. – Vol. 70. – P. 1126–1135.
6. Meta-Learning with Implicit Gradients / [A. Rajeswaran, C. Finn, S. Kakade et al.] // Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver, BC, Canada. – 2019. – P. 113–124.
7. Khabarlak K. Fast Facial Landmark Detection and Applications: A Survey [Electronic resource] / K. Khabarlak, L. Koriashkina // arXiv:2101.10808 [cs]. – 2021. – Access mode: https://arxiv.org/abs/2101.10808
8. Meta-learning with differentiable closed-form solvers [Electronic resource] / [L. Bertinetto, J. Henriques, P. Torr et al.] // 7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6–9, 2019. Access mode: https://openreview.net/forum?id=HyxnZh0ct7.
9. Snell J. Prototypical Networks for Few-shot Learning / J. Snell, K. Swersky, R.S. Zemel // Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4–9, 2017, Long Beach, CA, USA. – 2017. – P. 4077–4087.
10. Ravi S. Optimization as a Model for Few-Shot Learning [Electronic resource] / S. Ravi, H. Larochelle // 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings. – Access mode: https://openreview.net/forum?id=rJY0-Kcll.
11. Antoniou A. How to train your MAML [Electronic resource] / A. Antoniou, H. Edwards, A. J. Storkey // 7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6–9, 2019. Access mode: https://openreview.net/forum?id=HJGven05Y7.
12. Weng L. Meta-Learning: Learning to Learn Fast [Electronic resource] / L. Weng. – Access mode: https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html.
13. Yin W. Meta-learning for Few-shot Natural Language Processing: A Survey [Electronic resource] / W. Yin // CoRR. – 2020. – Vol. abs/2007.09604. – Access mode: https://arxiv.org/abs/2007.09604.
14. Generalizing from a Few Examples: A Survey on Few-shot Learning / [Y. Wang, Q. Yao, J. Kwok et al.] // ACM Comput. Surv. – 2020. – Vol. 53, № 3. – P. 63:1–63:34. DOI: 10.1145/3386252.
15. Guo Y. One-shot Face Recognition by Promoting Underrepresented Classes [Electronic resource] / Y. Guo, L. Zhang // CoRR. – 2017. – Vol. abs/1707.05574. – Access mode: http://arxiv.org/abs/1707.05574.
16. Koch G. Siamese neural networks for one-shot image recognition / G. Koch, R. Zemel, R. Salakhutdinov // ICML deep learning workshop. – Lille, 2015. – Vol. 2.
17. Matching Networks for One Shot Learning / [O. Vinyals, C. Blundell, T. Lillicrap et al.] // Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5–10, 2016, Barcelona, Spain. – 2016. – P. 3630–3638.
18. Meta-Learning with Memory-Augmented Neural Networks / [A. Santoro, S. Bartunov, M. Botvinick et al.] // Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19–24, 2016 : JMLR Workshop and Conference Proceedings. – JMLR.org, 2016. – Vol. 48. – P. 1842–1850.
19. Lake B.M. Human-level concept learning through probabilistic program induction / B.M. Lake, R. Salakhutdinov, J.B. Tenenbaum // Science. – 2015. – Vol. 350. – № 6266. – P. 1332–1338. – DOI: 10.1126/science.aab3050.
20. Nichol A. On First-Order Meta-Learning Algorithms [Electronic resource] / A. Nichol, J. Achiam, J. Schulman // CoRR. – 2018. – Vol. abs/1803.02999. – Access mode: http://arxiv.org/abs/1803.02999.
21. Meta-SGD: Learning to Learn Quickly for Few Shot Learning [Electronic resource] / [Z. Li, F. Zhou, F. Chen et al.] // CoRR. – 2017. – Vol. abs/1707.09835. – Access mode: http://arxiv.org/abs/1707.09835.
22. Zeiler M.D. Visualizing and Understanding Convolutional Networks / M.D. Zeiler, R. Fergus // Computer Vision – ECCV 2014 – 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part I : Lecture Notes in Computer Science. – Springer, 2014. – Vol. 8689. – P. 818–833. – DOI: 10.1007/978-3-319-10590-1_53.
23. Ioffe S. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift / S. Ioffe, C. Szegedy // Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6–11 July 2015: JMLR Workshop and Conference Proceedings. – JMLR.org, 2015. – Vol. 37. – P. 448–456.
24. Kingma D.P. Adam: A Method for Stochastic Optimization / D.P. Kingma, J. Ba // 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings. – 2015.
25. Krizhevsky A. Learning multiple layers of features from tiny images [Electronic resource] / A. Krizhevsky. – University of Toronto, 2009. – Access mode: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf

DOI 10.15588/1607-3274-2022-1-10

e-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2022. № 1
p-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2022. № 1

UDC 004.827:519.816

DEVELOPMENT OF MATHEMATICAL MODELS OF GROUP DECISION SYNTHESIS FOR STRUCTURING THE ROUGH DATA AND EXPERT KNOWLEDGE

Kovalenko I. I. – Dr. Sc., Professor, Professor of Department of Software Engineering, Petro Mohyla Black Sea National University, Mykolayiv, Ukraine.

Shved A. V. – Dr. Sc., Associate professor, Associate professor of Department of Software Engineering, Petro Mohyla Black Sea National University, Mykolayiv, Ukraine.

Davydenko Ye. O. – PhD, Associate professor, Head of Department of Software Engineering, Petro Mohyla Black Sea National University, Mykolayiv, Ukraine.

ABSTRACT

Context. The problem of aggregating the decision table attributes values formed out of group expert assessments as the classification problem was solved in the context of structurally rough set notation. The object of study is the process of the mathematical models synthesis for structuring and managing the expert knowledge that are formed and processed under incompleteness and inaccuracy (roughness).

Objective. The goal of the work is to develop a set of mathematical models for group expert assessments structuring for classification inaccuracy problem solving.

Method. A set of mathematical models for structuring the group expert assessments based on the methods of the theory of evidence has been proposed. This techniques allow to correctly manipulate the initial data formed under vagueness, imperfection, and inconsistency (conflict). The problems of synthesis of group decisions has been examined for two cases: taking into account decision table existing data, only, and involving additional information, i.e. subjective expert assessments, in the process of the aggregation of the experts' judgments.

Results. The outcomes gained can become a foundation for the methodology allowing to classify the groups of expert assessments with using the rough sets theory. This make it possible to form the structures modeling the relationship between the classification attributes of the evaluated objects, the values of which are formed out of the individual expert assessments and their belonging to the certain classes.

Conclusions. Models and methods of the synthesis of group decisions in context of structuring decision table data have been further developed. Three main tasks of structuring decision table data gained through the expert survey has been considered: the aggregation of expert judgments of the values of the decision attributes in the context of modeling of the relationship between the universe element and certain class; the aggregation of expert judgments of the values of the condition attributes; the synthesis of a group decision regarding the belonging of an object to a certain class, provided that the values of the condition attributes are also formed through the expert survey. The proposed techniques of structuring group expert assessments are the theoretical foundation for the synthesis of information technologies for the solution of the problems of the statistical and intellectual (classification, clustering, ranking and aggregation) data analysis in order to prepare the information and make the reasonable and effective decisions under incompleteness, uncertainty, inconsistency, inaccuracy and their possible combinations.

KEYWORDS: theory of evidence, rough set theory, aggregation, classification, inaccuracy, experts' judgments.

ABBREVIATIONS

bpa is a basic probability assignment;
DST is the Dempster-Shafer theory;
RST is the rough set theory;
DT is a decision table;
EP is an expert profile.

NOMENCLATURE

U is a non-empty, finite set of objects (the universe);
A is a set of primitive attributes;
C is a set of classification attributes;
E is a set of experts;
P is a set of profiles of expert preferences;
B is a set that reflects the judgments of all experts regarding the affiliation of the j-th object to a given class;
O is a set that reflects the preferences of all experts regarding the affiliation of the j-th object to a given class;
H is a set that reflects the judgments of all experts regarding the values of condition attributes for all objects;
Γ is a set of condition attributes values formed based on subjective and objective data for all objects;
B^{gr} is a set of profiles of group expert preferences in relation to the decision attributes values;
H^{gr} is a set of profiles of group expert preferences in relation to the condition attributes values;
$Bel()$ is a belief function of corresponding subset;
$Pl()$ is a plausibility function of corresponding subset;
B_i is a profile of the assessments of the i-th expert;
B_j^* is a set of expert judgments for the j-th object;
B_j^{**} is a set of unique expert judgments for the j-th object;
B_j^{comb} is a combined set of experts' evidences for j-th object by all experts;
M_j^* is a vector of mass functions (*bpa's*) formed on the judgments of all experts for j-th object;
M_j^{comb} is a vector of mass functions formed through the combination of corresponding mass functions by all experts for j-th object;

DOI 10.15588/1607-3274-2022-1-11

N is a maximum limit value of using scale;

R_j^* is a vector that contains a number of identical expert preferences regarding the belonging of the j-th object to a certain class;

$Y_k \subseteq b_j^i$ contains a number / name / marker of some class, to which j-th object was referred by the i-th expert in case when expert can refer the j-th object either to several classes or subgroups of classes with different degree of preference;

X_S is a subsets of the universe;

Z_k is a degree of preference of $Y_k \subseteq b_j^i$, $Y_k \succ a_q$ defined by the i-th expert for the j-th object;

$a_l^i(u_j)$ is a value of relevant condition attribute a_l formed by i-th expert in relation to j-th object;

a_q is a set of numbers, definitions, markers of the given classes;

$a_q^{gr}(u_j)$ is a group assessment regarding the belonging of the j-th object to a certain class;

b_j^i contains a number / name / marker of some class k_p, to which j-th object was referred by the i-th expert;

d is a total amount of the subsets (groups of elements) highlighted by the i-th expert for the j-th analyzed object;

d_j is the distance measure between its arguments;

k_p is a certain class, to which the j-th object was referred by the i-th expert;

$m()$ is a *bpa* of corresponding subset;

m_j^{**} is a vector that contains the values of *bpa's* of corresponding subsets;

n is a total number of experts;

o_j^i is the expert's subjective assessment (numeric value) proving that the element j-th object can be referred to a class k_p or a group of classes;

t is a total number of condition attributes;

z is a total number of the elements of the universe;

Ω is the frame of discernment;

θ_i is weighting coefficient (competence coefficient) of the i-th expert;

$[\pi]$ is some operator for processing the composite (group) expert assessment such as methods, algorithms;

2^{Ω} is a power-set of all possible subsets of Ω, including the empty set;

agr is an aggregation operator;

min is the function that gives the minimum value of its arguments;

$|\cdot|$ is a cardinality of the corresponding subset.

INTRODUCTION

The basic elements of the artificial intelligence systems such as pattern recognition systems, expert systems, decision support systems, etc. are knowledge bases formed out of such two approaches as object-oriented approach and object-structural approach [1, 2].

It is worth mentioning that, besides, the basic operation, which is realized as both above-mentioned approaches are used, is structuring the knowledge through their adjustment and classification as well as the typification of the highlighted classes. The currently mentioned procedures based on the generating specifications (functions) such as sum, difference, product, augmentation, etc., allow to form the families of subsets $\{X_1, X_2, ...X_n\}$ such that $X_i \subset X$, $X_i \neq \varnothing$, $X_i \cap X_j = \varnothing$ and $\cup X_i = \mathrm{X}$ $(i \neq j,\ i, j = \overline{1, n})$, which are based on the initial set of elements of knowledge X. Therefore, that allows to describe the knowledge by highlighting their properties and attributes or criteria. The currently described abstraction makes a base for choosing the basic concepts of knowledge processing such as production rules, predicate logic, and so on, as the artificial intelligence system design is done.

As a matter of fact, in the circumstances of real life, we quite often have to tackle the problems of getting the knowledge out of arrays of unstandardized, unprocessed, rough data and knowledge. The knowledge gained in the currently-mentioned process cannot be considered accurate, so it is not able to accurately classify them and to define a category of classification. Thus, it is connected, first of all, with the fact that the inflexibility of the existing models of knowledge presentation makes the analysts either unify or abridge the factual knowledge of the experts.

Thereupon, it is advisable to use the *RST*, the mathematical mechanism of which makes an inaccurate classification possible, which can be more factual than an accurate classification is, in practice [3]. Thus, according to the *RST*, a classification problem is formed as it is described beneath [4]. There are a set of multiple samples such as, for instance, a set of expert assessments of various types of objects, phenomena, events, and so on. Such initial set is called a learning set or universe. It is widely known that each sample belong to a class highlighted out of the given set of classes. Each sample possesses a typical set of classification attribute values. Taking that into account, the *RST* allows to model the relationship between the sample classified attribute values and sample membership in a certain class.

The object of study is the process of the mathematical models synthesis for structuring and managing the expert knowledge that are formed and processed under incompleteness and inaccuracy (roughness).

The subject of study is the models and methods of the group expert assessment analysis and structuring in the context of multi-alternative, incompleteness and inaccuracy (roughness).

The purpose of the work is to develop a set of mathematical models for group experts' assessments structuring for classification inaccuracy problem solving, based on the system application of methods of evidence theory and rough set theory.

1 PROBLEM STATEMENT

Assume that the given bounded set of analyzed objects (universe elements) is $U \neq \varnothing$. On the basis of U, it is pos-

DOI 10.15588/1607-3274-2022-1-11

sible to highlight the subset of universe elements X_S $X_S \subseteq U$ (a concept or category in U). Then, any family of concepts in U is considered to be abstract knowledge on U. Thus, the concepts form the division or classification of the currently-mentioned universe U. To put it in other words, in U, it is possible to highlight the family $C = \{X_s \mid s = \overline{1,n}\}$, whereas $X_s \subseteq U$, $X_s \neq \varnothing$, $X_s \cap X_t = \varnothing$ for $s \neq t$, $s,t = \overline{1,n}$. A family of classifications in U form a knowledge data base in U. Such knowledge data base is a set of aspects in the classification of universe objects.

Therefore, the existing knowledge system can be presented in a form of knowledge data base $K = (U, R)$, whereas $U = \{u_j \mid j = \overline{1,z}\}$ is a nonempty bounded set of elements (universe), R is equivalence relation, on the base of which one can form the equivalence classes (categories) of U-elements. Each category contains the elements possessing the common properties (attributes); within each category, the elements are considered indiscernible. The goal set $X_S \subseteq U$ is R-definite (R-accurate) if it is a unification of categories highlighted in U on the basis of R-relation. Otherwise, $X_S \subseteq U$ can be considered R-indefinite (R-inaccurate, R-rough).

The random knowledge data base K can correspond to the information system $S = (U, A, V, f)$, whereas $A = \{a_l \mid l = \overline{1,q}\}$ is a nonempty bounded set of primitive attributes; $V = \bigcup_{a_l \in A} V_{a_l}$, V_{a_l} is a set of attribute values a_l; $f \hat{} U \times A \to V$ is an data function such that $\forall a_l \in A, x \in U, f(x, a_l) \in V_{a_l}$.

To model the situation in which the element $u \in U$ can belong to a preliminarily-defined class based on the given set of attributes, the information system can be represented in a form of a DT $T = (U, A)$, whereas $A = C \cup D$ is a set of multiple condition attributes C ($|C| > 1$, $C = A \backslash \{a_q\}$ (classification attribute set) and a single-element subset D ($|D| = 1$, $D = \{a_q\}$) is a set of multiple decision attributes, the value of which describes the possible classes (a_q is a set of numbers, definitions, markers of the given classes), to which one can refer the elements of the initial universe. The relevant initial data of the information system and DT can be gained in different ways, i.e. both on the basis of objective and subjective initial information.

In the process of the analysis and partitioning the DT, during the group expertise, one can highlight the problems given below:

1) the problem of the aggregation of the appropriate values of the relevant decision attributes, i.e. the subjective expert values in relation to the values of $a_q(u_j)$, $a_q \in D$ formed out of the given set $a_l(u_j)$, $a_l \in C$, and synthesis of the group assessment in relation to the values of $a_q^{gr}(u_j)$, $u_j \in U$;

2) the problem of the aggregation of the appropriate values of the relevant condition attributes, i.e. the subjective expert values in relation to the values $a_l(u_j)$, $a_l \in C$, and synthesis of the group assessment in relation to the values of $a_l^{gr}(u_j)$, $u_j \in U$;

3) the problem of the group decision synthesis in relation to the membeship of the element $u_j \in U$ in the certain class: $u_j \to k_p$, $k_p \in a_q$ provided the relevant values of $a_l(u_j)$, ($a_l \in C$, $u_j \in U$) are also formed on the basis of the group expertise.

2 REVIEW OF THE LITERATURE

The *RST*, which was introduced by Z. Pawlak [5], allows to manipulate the initial data, which are considered rough as far as they are inaccurate and vague. The currently mentioned theory is for modeling the vagueness related with the universe elements belonging to the given goal set. To quantify such vagueness in [4, 5] measures of approximation accuracy and quality has been defined.

The theory is peculiar due to its mathematical mechanism helping to process the implicit arrays of unstandardized, i.e. inaccurate, rough, or unprocessed data and knowledge, and, thus, get the new knowledge. The theory is based on the fact that the knowledge is deeply involved in the human capability of classifying the subjects, phenomena, objects, situations, and so on, and so forth. They are reflected in the division (classification) of the relevant elements [3, 4, 5]. Such kind of division can be considered the knowledge presentation semantics. As a matter of fact, knowledge consists of the classification patterns of the application environment that is examined [3].

At the same time, knowledge is a kind of systematized information (objective or expert data) gained provided meeting the set criteria and structured for the solution of the set problem. In case if a sufficient amount of objective data, i.e. statistical, analytical, experimental, and empirical information, which can be gained my means of the methods of observation (registration), measurements (experiments, tests), is missing, it is advisable to involve a group of specialists (experts) in the certain application environment who form their judgments on the basis of the opinions and personal experience based on the interview, survey, focus-group with the methods of expert assessments. In such case, we can face a problem of aggregated expert assessment obtaining.

The analysis of a number of conventional methods of obtaining the expert assessments helps to arrive at the conclusion that different techniques of direct expert assessment averaging as well as the methods based on the various procedures of the comparison of the analyzed objects such as pairwise and multiple comparisons have become the most spread [6–9]. However, they are not deprived of a number of disadvantages. Obtaining the averaged assessment will be justified only if there is a high expert assessment consistency (proximity). In case if, there are several group supporting different opinions in the expert commission, it

DOI 10.15588/1607-3274-2022-1-11

will be no use simply averaging all the expert assessments. The main disadvantage of the methods based on the procedures of pairwise comparison is that they can be used for a small amount of compared elements. As the number of the latter grows, it is quite often difficult to achieve a high level of consistency of local priorities.

A favourable decision for the problems mentioned above should be provided by applying the advanced methods of the management of the indeterminacies, which have appeared in the last years. Therefore, we refer *DST*, i.e. evidence theory [10–12], and Theory of Plausible and Paradoxical Reasoning [13] to such methods. The mathematical apparatus of those theories allows to get the aggregated expert assessments using the technique of their combination. The choice of the combination rule depends on the study model (Dempster-Shafer model or Dezert-Smarandache model); the information on the conflicts between the expert evidence, which are combined; a structure of expert evidence; degree of consistency of expert evidences. In the works [13, 14], a number of recommendations for the choice of combination technique has been proposed.

3 MATERIALS AND METHODS

Let us consider the problem of aggregation of group expert assessments of decision attributes. Let us assume that the values of the *C*-subset elements are formed on the basis of the data gained through the objective studies based on the independent measurements, calculations and so on (objective data), and the values of the attribute a_q are formed out of the subjective data, i.e. data gained through the expert surveys.

Let, a group of experts $\mathrm{E} = \{E_i \mid i = \overline{1,n}\}$, taking into consideration the data of the given DT based on the values of the given set of *C*-tokens, formed the profiles of expert preferences $\mathrm{P} =< \mathrm{B} >$, whereas $\mathrm{B} = \{B_i \mid i = \overline{1,n}\}$. B_i-profile formed by the expert reflects its priorities in relation to the $u_j \in U$ ($j = \overline{1,z}$) element's membership in the given class $k_p \in a_q$ ($p = \overline{1,r}$, $r < z$). Thus, $B_i = \{b_j^i \mid j = \overline{1,z}\}$, whereas b_j^i contains a number / name / marker of some class $k_p \in a_q$, to which the object $u_j \in U$ was referred by the expert E_i.

The task is to synthesize the composite (group) profile $\mathrm{B}^{gr} = \{b_j^{gr} \mid j = \overline{1,z}\}$, $\underset{i=\overline{1,n}}{agr}(b_j^i) \to b_j^{gr}$, each b_j^{gr} element of which reflects a group solution, has a number, name or marker of some class $k_p \in a_q$, to which the object $u_j \in U$ was referred.

On the basis of the gained values of the B^{gr} composite profile for each object $u_j \in U$ that is examined, it will be able to set a class, to which it belongs: $\forall u_j \in U, j = \overline{1,z} : (u_j, b_j^{gr})$. The pair (u_j, k_p) sets the u_j object appurtenance to some class $k_p \in a_q$, the marker of which is preserved in b_j^{gr}.

A generalized scheme of the synthesis of the composite profile $\mathrm{B}^{gr} = \{b_j^{gr} \mid j = \overline{1,z}\}$ can be represented as follows, $i = \overline{1,n}$; $j = \overline{1,z}$:

$$B = \begin{pmatrix} B_1 \\ \dots \\ B_i \\ \dots \\ B_n \end{pmatrix} = \begin{pmatrix} b_1^1 & \dots & b_z^1 \\ \dots & \dots & \dots \\ b_1^i & \dots & b_z^i \\ \dots & \dots & \dots \\ b_1^n & \dots & b_z^n \end{pmatrix} \Rightarrow \begin{matrix} \\ \\ \\ \Uparrow \\ [\pi] \end{matrix} \begin{pmatrix} b_1^{gr} \\ \dots \\ b_j^{gr} \\ \dots \\ b_z^{gr} \end{pmatrix}^{-1} = \mathrm{B}^{gr}. \qquad (1)$$

For the synthesis of group (composite) expert assessments in modeling the relation "the element of the universe – the defined class", the mathematical notation of *DST* was used. While modeling the dependence "*U*-element – the *DT*-class", the following situations were studied:

1. $\forall u_j \in U$ whereas u_j element belongs to the only one class: $u_j \to k_p$;

2. $\exists u_j \in U$, which, according to the expert choice, can be referred to several classes: $u_j \to \{k_p, ..., k_s\}$, $p \neq s$, $p, s = \overline{1, r^*}$, $r^* < r$, $\forall p, s = \overline{1, r^*} : \{k_p \sim k_s\}$; in the result of modeling, $u_j \in U$ can be referred only to one class.

3. $\exists u_j \in U$, for which E_i cannot define a reference to any of the set classes: $u_j \to a_q$, $\forall p, s = \overline{1,r} : \{k_p \sim k_s\}$; as a result, $u_j \in U$ can be referred to the only one class.

The constraints, which are imposed on, and the conditions of the procedure of the expert survey can result in the following:

1. Using only the existing data and knowledge of the *DT* in the process of the expert assessment aggregation.

Let us examine the set a_q as the *DST* regards it. Let us assume, a_q is a frame of discernment, then, in the result of the expert survey, a system of subsets $B_i = \{b_j^i \mid j = \overline{1,z}\}$ will be formed, whereas b_j^i reflects E_i judgments in relation to the $u_j \in U$ -membership either in some $k_p \in a_q$ class, or in several classes (provided the expert defines a subgroup of classes, to one of which the $u_j \in U$ -object can be referred; the classes are equivalent inside the mentioned group). Thus, taking into account the *DST*-notation, b_j^i shall be regulated by a system of rules:

$$1.\ b_j^i = \{\varnothing\}; \qquad (2)$$

DOI 10.15588/1607-3274-2022-1-11

2. $\left|b_j^i\right| = 1$ – the expert has chosen and evaluated one element $k_p \in a_q$.

3. $\left|b_j^i\right| = h$, $h < \left|a_q\right|$ – the expert has highlighted h of the elements $k_p \in a_q$.

4. $b_j^i = a_q$ – it was difficult for the expert to assess / choose as far as all the elements of the set a_q are equivalent.

Aggregating the expert judgments is done according to the following suggested procedure:

1.1 Problem structuring. Let us highlight a subset of expert judgments $B_j^* = \{b_j^i\}$, $i = \overline{1,n}$ for each $u_j \in U$ and form a subset of the unique elements on the basis of those values: $B_j^{**} = \{b_t^*\}$, $t \le n$.

1.2 Define the vector $R_j^* = \{r_t^*\}$, whereas for $\forall t = \overline{1,|B_j^{**}|}$: $r_t^* = \text{count}(B_j^*(b_t^*))$ corresponds the number of the B_j^*-component that are equal to some value of $b_t^* \in B_j^{**}$.

1.3 Calculate the *bpa* masses for each subset B_j^{**}, taking into consideration the equation (formula):

$$m\{b_t^*\} = r_t^* \big/ |B_j^*|. \tag{3}$$

Thus, for each B_j^{**}-subset, it is possible to draw a vector $\mathrm{m}_j^{**} = \{m_t^* \mid t = \overline{1,|B_j^{**}|}\}$, the elements of which are in accord with the following constraints [10–12]:

$$0 \le m(X_j) \le 1, \quad m(\varnothing) = 0, \quad \sum_{X_j \in \Lambda} m(X_j) = 1, \tag{4}$$

whereas Λ corresponds to 2^{Ω}; $m : \Lambda \to [0,1]$.

1.4 The calculation of the upper and lower limits of the probability for each $k_p \in a_q$, which correspond the values of the belief function $Bel : \Lambda \to [0,1]$, [10–12]:

$$Bel(B) = \sum_{X_j \subseteq B,\, X_j \in \Lambda} m(X_j) \tag{5}$$

and plausibility function $Pl : \Lambda \to [0,1]$:

$$Pl(B) = \sum_{X_j \cap B \ne \varnothing,\, X_j \in \Lambda} m(X_j) \tag{6}$$

1.5 Forming the intervals $[Bel(\{k_p\}), Pl(\{k_p\})]$ for the subsets $k_p \in a_q$.

1.6 Choosing the optimal solution $b_{opt}^* \in a_q$ is done by means of the comparison of the intervals $[Bel(\{k_p\}), Pl(\{k_p\})]$, $\forall p = \overline{1,|a_q|}$ formed through the belief function and plausibility functions. The maximal interval, in which the lower value and upper value of the interval limits are the highest among the similar values of all the other intervals, corresponds the optimal solution: $b_{opt}^* = k_p : \max_p [Bel(\{k_p\}), Pl(\{k_p\})]$, $\forall p = \overline{1,r}$, $b_j^{gr} = b_{opt}^*$. Comparing all the embedded intervals, one can go from the interval values to crisp values. Thus, $b_j^{gr} = b_{opt}^*$, on the assumption that $b_{opt}^* \in a_q$.

2. Involving the additional information, i.e. subjective assessments, in the process of the expert judgment aggregation.

Situation 2.a. Expert E_i can refer the object $u_j \in U$ only either to a single class $k_p \in a_q$ or one subgroup of classes (the classes are considered equivalent within a highlighted subgroup, so the object $u_j \in U$ can be referred only to one of those classes).

Let us assume that, a group of experts $\mathrm{E} = \{E_i \mid i = \overline{1,n}\}$, based on the data of a given *DT*, constructed on the basis of the values of the organized set of *C*-tokens, formed the set of *EP's* $\mathrm{P} =< \mathrm{B}, \mathrm{O} >$. The P-set forms a tuple consisting of two components such as:

1) a set $\mathrm{B} = \{B_i \mid i = \overline{1,n}\}$, each element of which is $B_i = \{b_j^i \mid j = \overline{1,z}\}$ reflecting the reference mentioned by expert E_i regarding the affiliation of the element $u_j \in U$ ($j = \overline{1,z}$) either to a class $k_p \in a_q$ or several classes provided the expert can define a subgroup of classes, to one of which one can refer the object $u_j \in U$:

$$\forall u_j \in U \ \text{ expert } E_i : \begin{cases} u_j \to k_p \in a_q; \\ u_j \to \{k_p, \ldots, k_s\} \subseteq a_q. \end{cases} \tag{7}$$

2) a set $\mathrm{O} = \{O_i \mid i = \overline{1,n}\}$, each element of which is $O_i = \{o_j^i \mid j = \overline{1,z}\}$ reflecting the assessment of the E_i-expert's belief in the fact that the element $u_j \in U$ ($j = \overline{1,z}$) can be referred either to a class $k_p \in a_q$ or a subgroup of classes.

Thus, under the *DST* notation, b_j^i shall meet the standards of a system of rules (2); in its turn, o_j^i is expert's subjective assessment (probability) proving that the element $u_j \in U$ can be referred to a class $k_p \in a_q$ or a group of classes. The assessment of o_j^i can be repre-

DOI 10.15588/1607-3274-2022-1-11

sented within a set scale with using a range from 0 to N ($N > 0$). Under the assumption that $N \neq 1$, then the value o_j^i shall be normalized to a unit interval, i.e. $o_j^i \in [0;1]$.

Aggregating the expert judgments is done according to the following suggested procedure:

2.1 Problem structuring (partitioning). For each $u_j \in U$, let us highlight a set of expert judgements $B_j^* = \{b_j^i\}$, $i = \overline{1,n}$ and a set of assessments $O_j^* = \{o_j^i\}$, $i = \overline{1,n}$; let us form a subset of unique elements $B_j^{**} = \{b_t^*\}$, $t \le n$ of the $B_j^* = \{b_j^i\}$ on the basis of the obtained values.

2.2 According to the *DST*-notation, let us consider a set, i.e. the frame of discernment $\Omega = \{\omega_1, \omega_2\}$, whereas ω_1 for each E_i corresponds to the value of $b_j^i \in B_j^*$; $\omega_2 = a_q$ represents a complete lack of knowledge of the expert as to his choice. Under the assumption that $m(\omega_1)$ is the probability of the fact that the element $u_j \in U$ really belongs to the mentioned class, $m(\omega_1) = o_j^i$, in case of $o_j^i \in [0;1]$, $o_j^i \in O_j^*$; then the probability of the fact that the element can belong to some other class can be represented as $m(\omega_2) = 1 - m(\omega_1)$.

Thus, for each B_j^*, one will be able to get a set $\mathrm{M}_j^* = \{\mathrm{m}_j^i \mid i = \overline{1,n}\}$, whereas $\mathrm{m}_j^i = \{m(\omega_1), m(\omega_2)\}$ is a *bpa* vector of, as the expert E_i thinks, either right or wrong classification of the element $u_j \in U$, the elements of m_j^i satisfy (4).

2.3 Defining a procedure of the expert evidence aggregation (combination). For combining, one should choose a pair of expert evidences $b_j^i, b_j^h \in B_j^*$, such that under $i \neq h$: $\min d_J(\mathrm{m}_j^i, \mathrm{m}_j^h) \in [0;1]$ in compliance to one of the metric [15–18].

2.4 Aggregation of expert assessments is done through a combination of corresponding mass functions (*bpa's*) $\mathrm{M}_j^* = \{\mathrm{m}_j^i \mid i = \overline{1,n}\}$ and $B_j^* = \{b_j^i\}$, by all the experts E_i, $(i = \overline{1,n})$ for each $u_j \in U$ individually. In the result of the combination, a vector $B_j^{comb} = \{b_j^i \mid i = \overline{1,v}\}$, $v = 2^{\left|B_j^{**}\right|}$ and a vector $\mathrm{M}_j^{comb} = \{\mathrm{m}_j^i \mid i = \overline{1,v}\}$ can be obtained accordingly.

2.5. Calculation of the upper and lower bound of the plausibility for each $k_p \in a_q$ in compliance with (5) and (6) on the basis of the obtained B_j^{comb} and M_j^{comb}. Forming the intervals $[Bel(\{k_p\}), Pl(\{k_p\})]$ for the subsets $k_p \in a_q$.

2.6. The choice of an optimal $b_{opt}^* \in a_q$ is done through the comparison of the intervals $[Bel(\{k_p\}), Pl(\{k_p\})]$, $\forall p = \overline{1,\left|a_q\right|}$ formed with the belief and plausibility functions. The maximal interval corresponds the optimal solution. Thus, $b_j^{gr} = b_{opt}^*$ under the assumption that $b_{opt}^* \in a_q$.

Situation 2.b. Expert E_i can refer the object $u_j \in U$ either to several classes $k_p \in a_q$ or subgroups of classes with different degree of confidence (belief) in one's own choice. As far as the classes are considered equivalent within a highlighted subgroup, the object $u_j \in U$ can be referred only either to one class or a group of classes.

Let us assume that, analyzing the data of a given *DT*, constructed on the basis of the values of the organized set of *C*-tokens, a group of experts $\mathrm{E} = \{E_i \mid i = \overline{1,n}\}$ formed the set of *EP's* $\mathrm{P} = <\mathrm{B}, \mathrm{O}>$. A set of the *EP's* creates a tuple consisting of two components.

The first tuple component is $\mathrm{B} = \{B_i \mid i = \overline{1,n}\}$, each $B_i = \{b_j^i \mid j = \overline{1,z}\}$ element of which reflects the priorities mentioned by the expert E_i as to the membership of the element $u_j \in U$ ($j = \overline{1,z}$) in a class $k_p \in a_q$, or several classes. At the same time, $b_j^i = \{Y_k \mid k = \overline{1,d}\}$, $d \le 2^{\left|a_q\right|}$ is more than one value (several aimed classes or groups of classes). The second tuple component is $\mathrm{O} = \{O_i \mid i = \overline{1,n}\}$, each $O_i = \{o_j^i \mid j = \overline{1,z}\}$ element of which reflects the assessment of the E_i-expert belief in the fact that $u_j \in U$ ($j = \overline{1,z}$) is a member of the certain class $k_p \in a_q$ or a subgroup of classes. At the same time, $o_j^i = \{Z_k \mid k = \overline{1,d}\}$, $d \le 2^{\left|a_q\right|}$, $\forall i, j : \left|o_j^i\right| = \left|b_j^i\right|$, $i = \overline{1,n}$, $j = \overline{1,z}$.

Thus, taking into consideration, the *DST* notation, each element $Y_k \subseteq b_j^i$ shall meet the standards of the system of rules (2); in its turn, each element $Z_k \in o_j^i$ can create a probability, according to the expert's subjective assessment / belief), that the element $u_j \in U$ belongs to the certain class $k_p \in a_q$ or a group of classes. The assessment $Z_k \in o_j^i$ can be represented within the first given scale, using a range from 0 to the certain given N ($N > 0$).

Aggregation of the expert judgements is done in compliance with the suggested procedure, such as

2.1. Problem structuring (partitioning). Let us highlight a set of the expert's judgments $B_j^* = \{b_j^i\}$,

DOI 10.15588/1607-3274-2022-1-11

$i=\overline{1,n}$, and a set of expert's assessments $O_j^*=\{o_j^i\}$, $i=\overline{1,n}$ for each $u_j\in U$.

2.2 Defining the mass functions that correspond the highlighted subsets $Y_k\subseteq b_j^i$, $\forall b_j^i\in B_j^*$. For each formed system of subsets $b_j^i=\{Y_k\mid k=\overline{1,d}\}$, it will be possible to get a vector $\mathrm{m}_j^i=\{m_k\mid k=1,\overline{d+1}\}$, the elements of which correspond (4) and are calculated by the formulae, such as [10]:

$$m_k(Y_k)=\frac{Z_k\cdot\theta_i}{\sum\limits_{k=1}^{d}Z_k\cdot\theta_i+\sqrt{d}},$$
$$m_{d+1}(a_q)=\frac{\sqrt{d}}{\sum\limits_{k=1}^{d}Z_k\cdot\theta_i+\sqrt{d}}. \quad (8)$$

The value equaling $m_{d+1}(a_q)$ can reflect a degree of complete ignorance of E_i in relation to the membership of the object $u_j\in U$ in any class $k_p\in a_q$.

2.3. Defining the aggregation (combination) procedure of the expert judgments. For the combination, one can choose a pair of $b_j^i,b_j^h\in B_j^*$ such that under $i\neq h$: $\min d_J(\mathrm{m}_j^i,\mathrm{m}_j^h)\in[0;1]$ in accordance with one of metrics [15–18].

2.4. The aggregation of the *EP's* is done by the combination of the obtained *bpa's* $\mathrm{M}_j^*=\{\mathrm{m}_j^i\mid i=\overline{1,n}\}$ and $B_j^*=\{b_j^i\}$, by all the experts E_i, $(i=\overline{1,n})$, for each $u_j\in U$ individually, as well. The combination results are a vector $B_j^{comb}=\{Y_k^{comb}\mid k=\overline{1,v}\}$, $v\le 2^{|a_q|}$ and vector $\mathrm{M}_j^{comb}=\{\mathrm{m}(Y_k^{comb})\mid k=\overline{1,v}\}$, accordingly.

2.5. The calculation of the upper and lower bound for each $k_p\in a_q$ is done in compliance with (5) and (6), and on the basis of the obtained B_j^{comb} and M_j^{comb}, as well. The formation of the intervals $[Bel(\{k_p\}),Pl(\{k_p\})]$ for the subsets $k_p\in a_q$.

2.6. Choosing an optimal solution $b_{opt}^*\in a_q$ is done through the comparison of the intervals $[Bel(\{k_p\}),Pl(\{k_p\})]$, $\forall p=\overline{1,|a_q|}$. The maximal interval corresponds to the optimal solution. Thus, $b_j^{gr}=b_{opt}^*$ ($b_{opt}^*\in a_q$).

Let us consider the problem of aggregation of group expert assessments of condition attributes. For the *DT*, it is supposed that a *A*-set of primitive attributes is a union of two subsets $A=\{a_l\mid l=\overline{1,q-1}\}\cup a_q$, i.e. subsets of independent condition attributes $C=\{a_l\mid l=\overline{1,q-1}\}$ and one-element set of decision attribute $D=\{a_q\}$. Let us assume that, on the set *C,* it is possible to highlight a subset $C^*\subseteq C$, the elements of which are formed on the basis of the subjective data, i.e. data obtained by means of the expert survey. Let us enter the token $t=|C^*|$.

Let us assume that, examining the universe of discourse, a group of experts $\mathrm{E}=\{E_i\mid i=\overline{1,n}\}$ has formed the set of *EP's* such as $\mathrm{P}=<\mathrm{H}>$ or $\mathrm{P}=<\mathrm{H},\mathrm{O}>$, whereas $\mathrm{H}=\{H_i\mid i=\overline{1,n}\}$, $\mathrm{O}=\{O_i\mid i=\overline{1,n}\}$. Each element $H_i=\{H_j^i\mid j=\overline{1,z}\}$, $H_j^i=\{a_l^i(u_j)\mid l=\overline{1,t}\}$ of the first component of the E_i-profile reflects its preferences in relation to the values of the relevant condition attributes $a_l^i(u_j)$ of the element $u_j\in U$ ($j=\overline{1,z}$). The second component $\mathrm{O}=\{O_i\mid i=\overline{1,n}\}$ of the E_i-profile represents the expert's assessment of the belief in the correctness of his / her judgements, $O_i=\{O_j^i\mid j=\overline{1,z}\}$, $O_j^i=\{o_l^i(u_j)\mid l=\overline{1,t}\}$, whereas $o_l^i(u_j)$ is the assessment of the degree of confidence of the E_i in the set value of the attribute a_l for the element $u_j\in U$.

The task is to synthesize a group profile $\mathrm{H}^{gr}=\{H_j^{gr}\mid j=\overline{1,z}\}$, each element of which, i.e. $H_j^{gr}=\{a_l^{gr}(u_j)\mid l=\overline{1,t}\}$, represents a group solution and contains the aggregated values of the condition attributes $a_l^{gr}(u_j)$ of the $u_j\in U$ ($j=\overline{1,z}$), which are formed on the basis of the individual *EP's* $H_i=\{H_j^i\mid j=\overline{1,z}\}$, $\forall i=\overline{1,n}$, Fig. 1.

Taking into account the values of the H^{gr} composite profile, one can do the further examination and *DT*-data structuring.

The synthesis of the group decision is done according to the following procedure:

1. Problem structuring. Let a set of judgements $H_j^*=\{H_j^i\mid i=\overline{1,n}\}$, $\forall u_j\in U$ will be formed.

2. Aggregation of the group expert assessments $H_j^i\in H_j^*$.

The aggregation of the group expert's assessments is done individually for each attribute $a_l^i(u_j)$ by all the experts E_i, $i=\overline{1,n}$, i.e. $\forall l=\overline{1,t}:\underset{i}{agr}(a_l^i(u_j))\to H_j^{gr}$.

As an operator for processing the group expert's assessments of the relevant condition attributes can be one of the above schemes used.

DOI 10.15588/1607-3274-2022-1-11

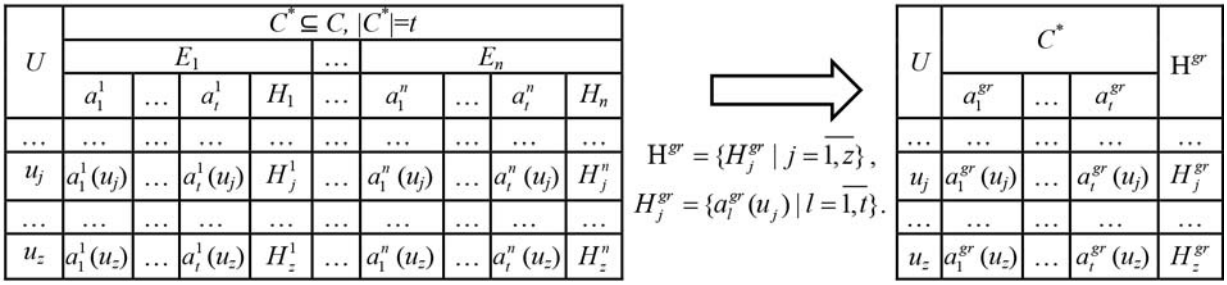


Figure 1 – The procedure for H^{gr} profile synthesis

The generalized scheme of the aggregation of H_j^i, $i = \overline{1,n}$, and construction of H_j^{gr} as a group expert assessment of the values of condition attributes a_l by the j-th object, can be represented in the following way:

$$H_j^* = \begin{pmatrix} a_1^1(u_j) & \dots & a_t^1(u_j) \\ \dots & \dots & \dots \\ a_1^i(u_j) & \dots & a_t^i(u_j) \\ \dots & \dots & \dots \\ a_1^n(u_j) & \dots & a_t^n(u_j) \end{pmatrix} \Rightarrow \begin{pmatrix} a_1^{gr}(u_j) \\ \dots \\ a_l^{gr}(u_j) \\ \dots \\ a_t^{gr}(u_j) \end{pmatrix}^{-1} = H_j^{gr}. \quad (9)$$

Similarly, the formation of the aggregated attribute values is made for each $u_j \in U$, $j = \overline{1,z}$.

Let us examine the problem of the group decision synthesis in relation to the membership of the element $u_j \in U$ in the given class provided the certain values of the relevant condition attributes of the $u_j \in U$ are also formed on the basis of the group expert evaluation.

In such a case, the expert evaluation shall be divided in two stages.

Stage 1. Solving the task of aggregation of the group expert assessments of condition attributes.

At that stage, a group of experts $\mathrm{E} = \{E_i \mid i = \overline{1,n}\}$ forms the set of *EP's* of $\mathrm{P} =< \mathrm{H} >$ or $\mathrm{P} =< \mathrm{H}, \mathrm{O} >$–types, whereas $\mathrm{H} = \{H_i \mid i = \overline{1,n}\}$, $\mathrm{O} = \{O_i \mid i = \overline{1,n}\}$. In the first case, the *EP* $H_i = \{H_j^i \mid j = \overline{1,z}\}$, $H_j^i = \{a_l^i(u_j) \mid l = \overline{1,t}\}$ formed by E_i reflects its preferences in relation to the values of the relevant condition attributes $a_l^i(u_j)$ of the $u_j \in U$ ($j = \overline{1,z}$). In the second case, the *EP* formed by the E_i contains an additional set $\mathrm{O} = \{O_i \mid i = \overline{1,n}\}$, $O_i = \{O_j^i \mid j = \overline{1,z}\}$, $O_j^i = \{o_l^i(u_j) \mid l = \overline{1,t}\}$ whereas $o_l^i(u_j)$ is an assessment of the degree of the E_i belief in the correctness of the fixed value of the attribute a_l for the $u_j \in U$.

The synthesis of a set of composite *EP's* $\mathrm{H}^{gr} = \{H_j^{gr} \mid j = \overline{1,z}\}$, each $H_j^{gr} = \{a_l^{gr}(u_j) \mid l = \overline{1,t}\}$ element of which reflects a group decision and contains the aggregated values of the relevant condition attributes $a_l^{gr}(u_j)$ of the $u_j \in U$ ($j = \overline{1,z}$) obtained on the basis of the individual *EP's* $H_i = \{a_l^i(u_j) \mid l = \overline{1,t}\}$, $\forall i = \overline{1,n}$, is carried out in accordance with the above-given scheme, i.e. a problem of aggregation of group expert assessments of condition attributes.

Stage 2. Solving the task of aggregation of the group expert assessments of decision attributes.

At the second stage, taking into account the values $\Gamma = \{\gamma_j \mid j = \overline{1,z}\}$, $\gamma_j = \{a_l(u_j) \mid l = \overline{1,|C|}\}$ of the given set of tokens (attributes) $A = \{a_l \mid l = \overline{1,q}\}$, $C = \{a_l \mid l = \overline{1,q-1}\}$, a group of experts $\mathrm{E} = \{E_i \mid i = \overline{1,n}\}$ forms the set of *EP's* $\mathrm{P} =< \mathrm{B} >$ whereas $\mathrm{B} = \{B_i \mid i = \overline{1,n}\}$.

We assume that each subset γ_j is formed on the basis of both:

1. The initial subjective data, i.e. under the group expert evaluation, the subjective values of the γ_j are formed out of the obtained at the first stage values of the relevant condition attributes $\mathrm{H}^{gr} = \{H_j^{gr} \mid j = \overline{1,z}\}$, $H_j^{gr} = \{a_l^{gr}(u_j) \mid l = \overline{1,t}\}$ such that $H_j^{gr} \subseteq \gamma_j$, $\forall j = \overline{1,z} : \left|H_j^{gr}\right| \le \left|\gamma_j\right|$.

2. The initial objective data. Under the assumption that $\forall j = \overline{1,z} : \left|H_j^{gr}\right| < \left|\gamma_j\right|$, the $\gamma_j = \{a_l(u_j)\}$, $\forall j = \overline{1,z} : \gamma_j \setminus H_j^{gr}$, are the values $a_l(u_j)$ formed on the basis of the objective data.

The E_i profile $B_i = \{b_j^i \mid j = \overline{1,z}\}$ represents its preferences in relation to the membership of the $u_j \in U$ ($j = \overline{1,z}$) in the given class $k_p \in a_q$ ($p = \overline{1,r}$, $r < z$), and the value b_j^i contains a number / name / or a marker of some class $k_p \in a_q$, to which the object $u_j \in U$ was referred by the E_i. The set of *EP's* can be represented in a form of $\mathrm{P} =< \mathrm{B}, \mathrm{O} >$. The second tuple component is a set $\mathrm{O} = \{O_i \mid i = \overline{1,n}\}$, each element of which, i.e.

DOI 10.15588/1607-3274-2022-1-11

$O_i = \{O_j^i \mid j = \overline{1,z}\}$, represents the degree of the E_i belief in the fact that the $u_j \in U$ ($j = \overline{1,z}$) is a member of either a certain class $k_p \in a_q$ or a subgroup of classes whereas $o_j^i = \{Z_k \mid k = \overline{1,d}\}$, $d \le 2^{|a_q|}$, $\forall i, j : |o_j^i| = |b_j^i|$, $i = \overline{1,n}$, $j = \overline{1,z}$.

The task is to synthesize the composite profile $\mathrm{B}^{gr} = \{b_j^{gr} \mid j = \overline{1,z}\}$ whereas b_j^{gr} represents a group decision in relation to the $u_j \in U$ membership in some $k_p \in a_q$ in accordance with the above-given scheme.

4 EXPERIMENTS

Let us demonstrate the above-suggested approaches, taking as a sample the solution of the problem of the group decisions synthesis in relation to the values of the relevant decision attributes. Let us assume that, taking into account the values of the formed set of tokens *C*, a group of experts $\mathrm{E} = \{E_i \mid i = \overline{1,5}\}$ evaluated the membership of the elements of the universe $u_j \in U$ ($j = \overline{1,3}$) in the given set of classes $a_q = \{k_p \mid p = \overline{1,3}\}$.

Sample 1. In the process of forming a group expert assessment, the only existing *DT* data and knowledge are used. The results of the expert survey are given in Table 1.

Table 1 – Expert profiles (Sample 1)

Objects	E_1	E_2	E_3	E_4	E_5
u_1	$\{k_2\}$	$\{k_2, k_3\}$	$\{k_2\}$	$\{k_1, k_3\}$	$\{k_3\}$
u_2	$\{k_1\}$	$\{k_2\}$	$\{k_1\}$	$\{k_1\}$	$\{k_1, k_2\}$
u_3	$\{k_2\}$	$\{k_1\}$	$\{k_1, k_2\}$	$\{k_3\}$	$\{k_2\}$

Sample 2. In the process of forming a group *EP's*, the embedded additional expert information is used; the expert E_i can refer the object $u_j \in U$ only to either one class $k_p \in a_q$ or a one subset of classes. The results of the expert survey are given in Table 2.

Sample 3. In the process of forming a group *EP's*, the embedded additional expert information is used; the expert E_i can refer the $u_j \in U$ either to several classes $k_p \in a_q$ or several subsets of classes with different degree of belief in one's own choice. The results of the expert survey are given in Table 3.

Table 2 – Expert profiles (Sample 2)

Objects	E_1		E_2		E_3		E_4		E_5	
	B_1	O_1	B_2	O_2	B_3	O_3	B_4	O_4	B_5	O_5
u_1	$\{k_2\}$	6	$\{k_2, k_3\}$	8	$\{k_2\}$	9	$\{k_1, k_3\}$	7	$\{k_3\}$	7
u_2	$\{k_1\}$	7	$\{k_2\}$	9	$\{k_1\}$	7	$\{k_1\}$	7	$\{k_1, k_2\}$	8
u_3	$\{k_2\}$	8	$\{k_1\}$	6	$\{k_1, k_2\}$	8	$\{k_3\}$	8	$\{k_2\}$	9

Table 3 – Expert profiles (Sample 3)

Objects	E_1		E_2		E_3		E_4		E_5	
	$Y_k \subseteq b_j^1$	$Z_k \subseteq o_j^1$	$Y_k \subseteq b_j^2$	$Z_k \subseteq o_j^2$	$Y_k \subseteq b_j^3$	$Z_k \subseteq o_j^3$	$Y_k \subseteq b_j^4$	$Z_k \subseteq o_j^4$	$Y_k \subseteq b_j^5$	$Z_k \subseteq o_j^5$
u_1	$\{k_2\}$	6	$\{k_2, k_3\}$	8	$\{k_2\}$	9	$\{k_1, k_3\}$	7	$\{k_3\}$	7
	$\{k_1\}$	7	$\{k_1\}$	5	–	–	$\{k_2\}$	5	$\{k_1\}$	9
	$\{k_3\}$	3	–	–	–	–	–	–	–	–
u_2	$\{k_1\}$	8	$\{k_2\}$	9	$\{k_1\}$	7	$\{k_1\}$	7	$\{k_1, k_2\}$	8
	$\{k_2, k_3\}$	5	–	–	$\{k_3\}$	4	$\{k_2\}$	9	–	–
u_3	$\{k_2\}$	8	$\{k_1\}$	6	$\{k_1, k_2\}$	8	$\{k_3\}$	8	$\{k_2\}$	9

Tables 1–3 represents only the subjective judgments made by five experts as to the membership of the elements of the given universe in a fixed set of classes. In such a case, the values of the classified attributes of the universe elements are omitted on purpose as far as they do not matter, by any means, for the problem that is examined.

The elements of the set O_i (Table 2) and set Z_k (Table 3) were evaluated according to the ten-point scale (zero stands for the lowest degree of preference and ten stands for an absolute degree of preference).

5 RESULTS

Let us examine the practical realization of the above methods for synthesizing a group decision in relation to the membership of the element $u_1 \in U$ in a class $k_p \in a_q$.

In the process of analyzing the data from Table 1, it can be seen that, taking into account $a_q = \{k_p \mid p = \overline{1,3}\}$, for $u_1 \in U$, a group of experts formed a set $B_1^* = \{\{k_2\}, \{k_2, k_3\}, \{k_2\}, \{k_1, k_3\}, \{k_3\}\}$ on the basis of which it will be possible to form a set $B_1^{**} = \{\{k_2\}, \{k_3\}, \{k_1, k_3\}, \{k_2, k_3\}\}$, and a vector $R_1^* = \{2,1,1,1\}$.

Let us calculate the basic probability assignment for each element of the set B_1^{**} according to equation (3):

DOI 10.15588/1607-3274-2022-1-11

$m\{k_2\} = 2/5$; $m\{k_3\} = 1/5$;
$m\{k_1, k_3\} = 1/5$; $m\{k_2, k_3\} = 1/5$.

Let us calculate the value of the functions (5) and (6) for each element of the set a_q:

$$k_1 : \begin{cases} Bel(\{k_1\}) = m(\{k_1\}) = 0; \\ Pl(\{k_1\}) = 0.2; \end{cases}$$
$$k_2 : \begin{cases} Bel(\{k_2\}) = m(\{k_2\}) = 0.4; \\ Pl(\{k_2\}) = 0.6; \end{cases}$$
$$k_3 : \begin{cases} Bel(\{k_3\}) = m(\{k_3\}) = 0.2; \\ Pl(\{k_3\}) = 0.6. \end{cases}$$

After taking a look at the above calculations, one can see that the $b^*_{opt} = \{k_2\}$ is an optimal one. Thus, we obtain $u_1 \to k_2$ and $b_1^{gr} = \{k_2\}$, accordingly.

In the process of analysis the data given in Table 2, one can see that, taking into account $a_q = \{k_p \mid p = \overline{1,3}\}$, for $u_1 \in U$, the experts formed a set $B_1^* = \{\{k_2\}, \{k_2, k_3\}, \{k_2\}, \{k_1, k_3\}, \{k_3\}\}$. On the basis of the values of the latter, we can form a set $B_1^{**} = \{\{k_2\}, \{k_3\}, \{k_1, k_3\}, \{k_2, k_3\}\}$ and a set $O_1^* = \{6, 8, 9, 7, 7\}$, as well.

The *bpa's* of the formed focal elements are given in Table 4.

Let us calculate the combined values of the *bpa's* of the highlighted subsets:

$m\{k_2\} = 0.44$; $m\{k_3\} = 0.47$;
$m\{k_2, k_3\} = 0.003$; $m\{k_1, k_3\} = 0.0863$.
$m\{k_1, k_2, k_3\} = 0.0007$;

Let us calculate the values of the functions (5) and (6) for each element of the set a_q:

$$k_1 : \begin{cases} Bel(\{k_1\}) = 0; \\ Pl(\{k_1\}) = 0.087; \end{cases}$$
$$k_2 : \begin{cases} Bel(\{k_2\}) = 0.44; \\ Pl(\{k_2\}) = 0.444; \end{cases}$$
$$k_3 : \begin{cases} Bel(\{k_3\}) = 0.47; \\ Pl(\{k_3\}) = 0.56. \end{cases}$$

Taking into account the above calculations, one can see that the $b^*_{opt} = \{k_3\}$ is an optimum choice. Thus, we obtain $u_1 \to k_3$ and $b_1^{gr} = \{k_3\}$, accordingly.

In the process of the analysis of the data from Table 3, one can see that the experts formed a set $B_1^* = \{b_1^i\}$ and a set of assessments $O_1^* = \{o_1^i\}$, $i = \overline{1,n}$ for $u_1 \in U$, on the basis of $a_q = \{k_p \mid p = \overline{1,3}\}$, whereas

$b_1^1 = \{\{k_1\}, \{k_2\}, \{k_3\}\}$; $o_1^1 = \{7, 6, 3\}$;
$b_1^2 = \{\{k_1\}, \{k_2, k_3\}\}$; $o_1^2 = \{5, 8\}$;
$b_1^3 = \{\{k_2\}\}$; $o_1^3 = \{9\}$;
$b_1^4 = \{\{k_2\}, \{k_1, k_3\}\}$; $o_1^4 = \{5, 7\}$;
$b_1^5 = \{\{k_1\}, \{k_3\}\}$; $o_1^5 = \{9, 7\}$.

The *bpa's* of the formed focal elements are given in Table 5.

Table 4 – The *bpa's* of the formed focal elements (Sample 2)

Objects	E_1		E_2		E_3		E_4		E_5	
	$m(\omega_1)$	$m(\omega_2)$	$m(\omega_1)$	$m(\omega_2)$	$m(\omega_1)$	$m(\omega_2)$	$m(\omega_1)$	$m(\omega_2)$	$m(\omega_1)$	$m(\omega_2)$
u_1	0.6	0.4	0.8	0.2	0.9	0.1	0.7	0.3	0.7	0.3
u_2	0.7	0.3	0.9	0.1	0.7	0.3	0.7	0.3	0.8	0.2
u_3	0.8	0.2	0.6	0.4	0.8	0.2	0.8	0.2	0.9	0.1

Table 5 – The *bpa's* of the formed focal elements (Sample 3)

Objects	E_1		E_2		E_3		E_4		E_5	
	$Y_k \subseteq b_j^1$	$m(Y_k)$	$Y_k \subseteq b_j^2$	$m(Y_k)$	$Y_k \subseteq b_j^3$	$m(Y_k)$	$Y_k \subseteq b_j^4$	$m(Y_k)$	$Y_k \subseteq b_j^5$	$m(Y_k)$
u_1	$\{k_2\}$	0.34	$\{k_2, k_3\}$	0.55	$\{k_2\}$	0.90	$\{k_1, k_3\}$	0.52	$\{k_3\}$	0.40
	$\{k_1\}$	0.39	$\{k_1\}$	0.35	$\{k_1, k_2, k_3\}$	0.10	$\{k_2\}$	0.37	$\{k_1\}$	0.52
	$\{k_3\}$	0.17	$\{k_1, k_2, k_3\}$	0.10	–	–	$\{k_1, k_2, k_3\}$	0.11	$\{k_1, k_2, k_3\}$	0.08
	$\{k_1, k_2, k_3\}$	0.10	–	–	–	–	–	–	–	–
u_2	$\{k_1\}$	0.55	$\{k_2\}$	0.9	$\{k_1\}$	0.56	$\{k_1\}$	0.40	$\{k_1, k_2\}$	0.89
	$\{k_2, k_3\}$	0.35	$\{k_1, k_2, k_3\}$	0.1	$\{k_3\}$	0.32	$\{k_2\}$	0.52	$\{k_1, k_2, k_3\}$	0.11
	$\{k_1, k_2, k_3\}$	0.10	–	–	$\{k_1, k_2, k_3\}$	0.12	$\{k_1, k_2, k_3\}$	0.08	–	–
u_3	$\{k_2\}$	0.89	$\{k_1\}$	0.86	$\{k_1, k_2\}$	0.89	$\{k_3\}$	0.89	$\{k_2\}$	0.90
	$\{k_1, k_2, k_3\}$	0.11	$\{k_1, k_2, k_3\}$	0.14	$\{k_1, k_2, k_3\}$	0.11	$\{k_1, k_2, k_3\}$	0.11	$\{k_1, k_2, k_3\}$	0.10

DOI 10.15588/1607-3274-2022-1-11

Let us calculate the combined values of the *bpa's* of highlighted subsets in compliance with (8):

$$m\{k_1\} = 0.331; \qquad m\{k_2\} = 0.438;$$
$$m\{k_3\} = 0.217; \qquad m\{k_1, k_3\} = 0.013;$$
$$m\{k_2, k_3\} = 0.0002; \qquad m\{k_1, k_2, k_3\} = 0.0008.$$

Let us calculate values of the functions (5) and (6) for each element of the set a_q:

$$k_1 : \begin{cases} Bel(\{k_1\}) = 0.331; \\ Pl(\{k_1\}) = 0.3448. \end{cases}$$

$$k_2 : \begin{cases} Bel(\{k_2\}) = 0.438; \\ Pl(\{k_2\}) = 0.439. \end{cases}$$

$$k_3 : \begin{cases} Bel(\{k_3\}) = 0.217; \\ Pl(\{k_3\}) = 0.231. \end{cases}$$

Taking a look at the estimations described above, one can see that the $b^*_{opt} = \{k_2\}$ is an optimum. Thus, we obtain $u_1 \to k_2$ and $b_1^{gr} = \{k_2\}$, accordingly.

6 DISCUSSION

The problems of the group decisions synthesis while modeling the relationship between the element of universe and definite class either in case if one takes into consideration only the existing *DT* data or if the additional information, i.e. subjective expert assessments, is involved in the process of aggregating the expert judgment, have been studied. In solving the problem of the aggregation of the relevant *DT* attributes formed on the basis of the subjective assessments of the expert group, the situations are considered when an expert can either refer a universe object only to one class, i.e. a single subgroup of classes when the classes are considered equivalent within a highlighted subset, or define that a universe object can refer to several separate classes, i.e. subsets of classes, with different degree of confidence in one's own choice, i.e. the classes can be considered equivalent within a highlighted subset.

To make the aggregated expert assessments, a mathematical mechanism of evidence theory has been used. Unlike the existing techniques of the expert evidence aggregation, that allowed to synthesize the group decisions in the context of multiple alternatives, incompleteness, inaccuracy and inconsistency (conflict), as well as to model the uncertainty and not to strictly restrain the expert in his / her personal choice. To put it in other words, the expert can choose not only a single priority but also can form the clustered ranges of objects, setting a degree of confidence in his / her own choice.

CONCLUSIONS

The problems, which arise in the process of the analysis and structuring of *DT* data in the context of the group expert evaluation have been formulated. A set of mathematical models for structuring the *DT* data obtained out of the expert assessments, which are formed and processed in the context of inaccuracy (roughness), imperfection, and multiple alternatives in the process of solving the inaccurate classification problem, has been proposed.

The scientific novelty of obtained results is that the models and methods of synthesizing the group decisions and *DT* data structuring are received the further development. The next problems of synthesizing the group decisions and *DT* data structuring have been solved: the synthesis of the group assessments of the values of the relevant decision attributes, the synthesis of the group assessments of the values of the relevant condition attributes, and the synthesis of the group assessments concerning the membership of the universe object in the given class provided the appropriate values of the relevant condition attributes of the object are also formed on the basis of the group expert evaluation. The suggested techniques are based on the mathematical notation of the evidence theory. That allowed processing the group expert assessments under vagueness, imperfection, and inconsistency (conflict).

The practical significance of the obtained results implies that the suggested techniques form a theoretical substratum for plotting the methods, algorithms, and information technologies for intelligent support of the decision-making process, and its implementing in the automated decision-support systems for an inaccurate classification problem solving. The obtained results can be helpful in the formation of the bases of the expert's knowledge in different universes of discourse.

The prospects for further research imply the development of the procedure of reducing the *DT* knowledge formed on the basis of the individual expert judgments, especially in the context of the incomplete expert data.

ACKNOWLEDGEMENTS

The work is partially supported by the state research project of Petro Mohyla Black Sea National University "Development of modern information and communication technologies for the management of intellectual resources to decision-making support of operational management" (research project no. 0121U107831, financed by the Government of Ukraine).

REFERENCES

1. Jakus G., Milutinovic V., Omerovic S., Tomazic S. Concepts, ontologies, and knowledge representation. New York, Springer, 2013, 73 p. DOI: 10.1007/978-1-4614-7822-5
2. Patel-Schneider P. F. Practical, object-based knowledge representation for knowledge-based systems, *Information Systems*, 1990, Vol. 15(1), pp. 9–19. DOI: 10.1016/0306-4379(90)90013-F.
3. Uzga-Rebrovs O. Nenoteiktibu parvaldisana. Rezekne, RA Izdevnieciba, 2010, Vol. 3, 560 p.
4. Pawlak Z. Rough sets, *International Journal of Computer & Information Sciences,* 1982, Vol. 11(5), pp. 341–356. DOI: 10.1007/BF01001956
5. Pawlak Z. Rough sets, theoretical aspects of reasoning about data. Boston, Kluwer Academic Publishers, 1991, 229 p.
6. Alinezhad A., Khalili J. New methods and applications in multiple attribute decision making (MADM). Cham, Springer, 2019, 257 p. DOI: 10.1007/978-3-030-15009-9

DOI 10.15588/1607-3274-2022-1-11

7. Ishizaka A., Nemery P. Multicriteria decision analysis: methods and software. New York, John Wiley & Sons, 2013, 312 p. DOI: 10.1002/9781118644898
8. Radford K. J. Individual and small group decisions. New York, Springer, 1989, 175 p. DOI: 10.1007/978-1-4757-2068-6
9. Saaty T. The Analytic Hierarchy Process: panning, priority setting, resource allocation. Front cover. New York, McGraw Hill, 1980, 287 p.
10. Beynon M. J., Curry B., Morgan P. The Dempster-Shafer theory of evidence: an alternative approach to multicriteria decision modeling, *Omega*, 2000, Vol. 28, No. 1, pp. 37–50. DOI: 10.1016/S0305-0483(99)00033-X
11. Dempster A. P. Upper and lower probabilities induced by a multi-valued mapping, *Annals of Mathematical Statistics*, 1967, Vol. 38(2), pp. 325–339. DOI: 10.1214/aoms/1177698950
12. Shafer G. A mathematical theory of evidence. Princeton, Princeton University Press, 1976, 297 p.
13. Smarandache F., Dezert J. Advances and applications of DSmT for information fusion. Rehoboth, American Research Press, 2004, Vol. 1, 760 p.
14. Sentz K., Ferson S. Combination of evidence in Dempster-Shafer theory. Technical report SAND 2002-0835. Albuquerque, Sandia National Laboratories, 2002, 94 p.
15. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distribution, *Bulletin of the Calcutta Mathematical Society*, 1943, Vol. 35, pp. 99–110.
16. Cuzzolin F. A geometric approach to the theory of evidence, *Transactions on Systems, Man, and Cybernetics*, Part C: Applications and Reviews, 2007, Vol. 38(4), pp. 522–534. DOI: 10.1109/TSMCC.2008.919174
17. Jousselme A. L., Grenier D., Boss´e E. A new distance between two bodies of evidence, *Information Fusion,* 2001, Vol. 2, pp. 91–101. DOI: 10.1016/S1566-2535(01)00026-4
18. Tessem B. Approximations for efficient computation in the theory of evidence, *Artificial Intelligence*, 1993, Vol. 61, pp. 315–329. DOI: 10.1016/0004-3702(93)90072-J

Received 08.11.2021.
Accepted 18.12.2021.

УДК 004.827:519.816

РОЗРОБКА МАТЕМАТИЧНИХ МОДЕЛЕЙ СИНТЕЗУ ГРУПОВИХ РІШЕНЬ СТРУКТУРИЗАЦІЇ ГРУБИХ ДАНИХ ТА ЕКСПЕРТНИХ ЗНАНЬ

Коваленко І. І. – д-р техн. наук, професор, професор кафедри інженерії програмного забезпечення Чорноморського національного університету імені Петра Могили, Миколаїв, Україна.

Швед А. В. – д-р техн. наук, доцент, доцент кафедри інженерії програмного забезпечення Чорноморського національного університету імені Петра Могили, Миколаїв, Україна.

Давиденко Є. О. – канд. техн. наук, доцент, завідувач кафедри інженерії програмного забезпечення Чорноморського національного університету імені Петра Могили, Миколаїв, Україна.

АНОТАЦІЯ

Актуальність. Розглянуті питання агрегування значень атрибутів таблиці рішень, сформованих на основі групових експертних оцінок при вирішенні задачі неточної класифікації в рамках нотації теорії грубих множин. Об’єктом дослідження є процеси синтезу математичних моделей структуризації та управління експертними знаннями, які формуються та обробляються в умовах неточності (грубості) та неповноти. Мета роботи – розробка математичних моделей структуризації групових експертних оцінок при вирішенні задачі «неточної класифікації».

Метод. Запропоновано комплекс математичних моделей структуризації групових експертних оцінок, в основу яких покладено методи теорії свідоцтв, які дозволяють коректно оперувати з вихідними даними, сформованими в умовах невизначеності, неповноти, неузгодженості (конфлікту). Розглянуті питання синтезу групових рішень для двох випадків: тільки на основі існуючих даних таблиці рішень, і з залученням додаткової інформації (суб’єктивних експертних оцінок) в процесі агрегування суджень експертів.

Результати. Отримані результати можуть бути покладені в основу методики, що дозволяє виконувати класифікацію групових експертних оцінок із застосуванням теорії грубих множин. Це дає можливість формувати структури, що моделюють залежність між класифікаційними атрибутами оцінюваних об’єктів, значення яких формуються на основі індивідуальних експертних оцінок, і їх приналежністю відповідним класам.

Висновки. Дістали подальшого розвитку моделі та методи синтезу групових рішень у контексті структурування даних таблиці рішень. Три основні задачі структурування даних таблиці рішень, одержаних у результаті експертного опитування, було розглянуто: агрегування експертних суджень щодо значень атрибутів рішень при моделюванні залежності «елемент універсуму – визначений клас»; агрегування експертних оцінок щодо значень атрибутів умов; синтез групового рішення щодо належності об'єкта до певного класу за умови, що значення атрибутів умов також формуються шляхом експертного опитування. Запропоновані техніки структуризації групових експертних оцінок становлять теоретичне підґрунтя для синтезу інформаційних технологій вирішення задач статистичного та інтелектуального (класифікація, кластеризація, ранжування, агрегування) аналізу даних з метою підготовки інформації для прийняття обґрунтованих та ефективних рішень в умовах неповноти, невизначеності, неузгодженості неточності та їх можливих комбінацій.

КЛЮЧОВІ СЛОВА: теорія свідоцтв, теорія грубих множин, агрегування, класифікація, неточність, експертні оцінки.

УДК 004.827:519.816

РАЗРАБОТКА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ СИНТЕЗА ГРУППОВЫХ РЕШЕНИЙ СТРУКТУРИЗАЦИИ ГРУБЫХ ДАННЫХ И ЭКСПЕРТНЫХ ЗНАНИЙ

Коваленко И. И. – д-р техн. наук, профессор, профессор кафедры инженерии программного обеспечения Черноморского национального университета имени Петра Могилы, Николаев, Украина.

DOI 10.15588/1607-3274-2022-1-11

Швед А. В. – д-р техн. наук, доцент, доцент кафедры инженерии программного обеспечения Черноморского национального университета имени Петра Могилы, Николаев, Украина.

Давыденко Е. А. – канд. техн. наук, доцент, заведующий кафедрой инженерии программного обеспечения Черноморского национального университета имени Петра Могилы, Николаев, Украина.

АННОТАЦИЯ

Актуальность. Рассмотрены вопросы агрегирования значений атрибутов таблицы решений, сформированных на основе групповых экспертных оценок при решении задач неточной классификации в рамках нотации теории грубых множеств. Объектом исследования являлись процессы синтеза математических моделей структуризации и управления экспертными знаниями, которые формируются и обрабатываются в условиях неточности (грубости) и неполноты. Цель работы – разработка математических моделей структуризации групповых экспертных оценок при решении задачи «неточной классификации».

Метод. Предложено комплекс математических моделей структуризации групповых экспертных оценок, в основу которых положены методы теории свидетельств, позволяющие корректно оперировать с исходными данными, сформированными в условиях неопределенности, неполноты, несогласованности (конфликта). Рассмотрены вопросы синтеза групповых решений для двух случаев: только на основе существующих данных таблицы решений, и с привлечением дополнительной информации (субъективных экспертных оценок) в процессе агрегирования суждений экспертов.

Результаты. Полученные результаты могут быть положены в основу методики, позволяющей выполнять классификацию групповых экспертных оценок с применением теории грубых множеств. Это позволяет формировать структуры, моделирующие зависимость между классификационными атрибутами оцениваемых объектов, значения которых формируются на основе индивидуальных экспертных оценок, и их принадлежностью соответствующим классам.

Выводы. Получили дальнейшее развитие модели и методы синтеза групповых решений в контексте структурирования данных таблицы решений. Три основные задачи структурирования данных таблицы решений, полученных в результате экспертного опроса, были рассмотрены: агрегирование экспертных оценок значений атрибутов решений при моделировании зависимости «элемент универсума – определенный класс»; агрегирование экспертных оценок значений атрибутов условий; синтез группового решения о принадлежности объекта к некоторому классу при условии, что значения атрибутов условий также формируются за результатами экспертного опроса. Предложенные техники структуризации групповых экспертных оценок составляют теоретическую основу для синтеза информационных технологий решения задач статистического и интеллектуального (классификация, кластеризация, ранжирование, агрегирование) анализа данных с целью подготовки информации для принятия обоснованных и эффективных решений в условиях неполноты, неопределенности, несогласованности неточности их возможных комбинаций.

КЛЮЧЕВЫЕ СЛОВА: теория свидетельств, теория грубых множеств, агрегирование, классификация, неточность, экспертные оценки.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Jakus G. Concepts, ontologies, and knowledge representation / G. Jakus, V. Milutinovic, S. Omerovic, S. Tomazic. – New York: Springer, 2013. – 73 p. DOI: 10.1007/978-1-4614-7822-5
2. Patel-Schneider P. F. Practical, object-based knowledge representation for knowledge-based systems / P. F. Patel-Schneider // Information Systems. – 1990. – Vol. 15(1). – P. 9–19. DOI: 10.1016/0306-4379(90)90013-F.
3. Uzga-Rebrovs O. Nenoteiktibu parvaldisana / O. Uzga-Rebrovs. – Rezekne: RA Izdevnieciba, 2010. – Vol. 3. – 560 lpp.
4. Pawlak Z. Rough sets / Z. Pawlak // International Journal of Computer & Information Sciences. – 1982. – Vol. 11(5). – P. 341–356. DOI: 10.1007/BF01001956
5. Pawlak Z. Rough sets, theoretical aspects of reasoning about data / Z. Pawlak. – Boston: Kluwer Academic Publishers, 1991. – 229 p.
6. Alinezhad A. New methods and applications in multiple attribute decision making (MADM) / A. Alinezhad, J. Khalili. – Cham : Springer, 2019. – 257 p. DOI: 10.1007/978-3-030-15009-9
7. Ishizaka A. Multicriteria decision analysis: methods and software / A. Ishizaka, P. Nemery. – New York; John Wiley & Sons, 2013. – 312 p. DOI: 10.1002/9781118644898
8. Radford K. J. Individual and small group decisions / K. J. Radford. – New York: Springer, 1989. – 175 p. DOI: 10.1007/978-1-4757-2068-6
9. Saaty T. The Analytic Hierarchy Process: panning, priority setting, resource allocation. Front cover. / T. Saaty. – New York: McGraw Hill, 1980. – 287 p.
10. Beynon M. J. The Dempster-Shafer theory of evidence: an alternative approach to multicriteria decision modeling / M. J. Beynon, B. Curry, P. Morgan // Omega. – 2000. – Vol. 28, № 1. – P. 37–50. DOI: 10.1016/S0305-0483(99)00033-X
11. Dempster A. P. Upper and lower probabilities induced by a multi-valued mapping / A. P. Dempster // Annals of Mathematical Statistics. – 1967. – Vol. 38(2). – P. 325–339. DOI: 10.1214/aoms/1177698950
12. Shafer G. A mathematical theory of evidence / G. Shafer. – Princeton: Princeton University Press, 1976. – 297 p.
13. Smarandache F. Advances and applications of DSmT for information fusion / F. Smarandache, J. Dezert. – Vol. 1. – Rehoboth: American Research Press, 2004. – 760 p.
14. Sentz K. Combination of evidence in Dempster-Shafer theory. Technical report SAND 2002-0835 / K. Sentz, S. Ferson. – Albuquerque: Sandia National Laboratories, 2002. – 94 p.
15. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distribution / A. Bhattacharyya // Bulletin of the Calcutta Mathematical Society. – 1943. – Vol. 35. – P. 99–110.
16. Cuzzolin F. A geometric approach to the theory of evidence / F. Cuzzolin // Transactions on Systems, Man, and Cybernetics (Part C: Applications and Reviews). – 2007. – Vol. 38(4). – P. 522–534. DOI: 10.1109/TSMCC.2008.919174
17. Jousselme A. L. A new distance between two bodies of evidence / A. L. Jousselme, D. Grenier, E. Boss´e // Information Fusion. – 2001. – Vol. 2. – P. 91–101. DOI: 10.1016/S1566-2535(01)00026-4
18. Tessem B. Approximations for efficient computation in the theory of evidence / B. Tessem // Artificial Intelligence. – 1993. – Vol. 61. – P. 315–329. DOI: 10.1016/0004-3702(93)90072-J

DOI 10.15588/1607-3274-2022-1-11

e-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2022. № 1
p-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2022. № 1

UDC 004.94

DEVELOPING A FUZZY RISK ASSESSMENT MODEL FOR ERP-SYSTEMS

Kozhukhivskyi A. D. – Dr. Sc., Professor, Professor Department of Information and Cybernetic security of State University of Telecommunications, Kyiv, Ukraine.

Kozhukhivska O. A. – Dr. Sc., Associate Professor Department of Information and Cybernetic security of State University of Telecommunications, Kyiv, Ukraine.

ABSTRACT

Context. Because assessing information security risks is a complex and complete uncertainty process, and uncer-tainties are a major factor influencing valuation performance, it is advisable to use fuzzy methods and models that are adaptive to non-calculated data. The formation of vague assessments of risk factors is subjective, and risk assessment depends on the practical results obtained in the process of processing the risks of threats that have already arisen during the functioning of the organization and experience of information security professionals. Therefore, it will be advisable to use models that can adequately assess fuzzy factors and have the ability to adjust their impact on risk assessment. The greatest performance indicators for solving such problems are neuro-fuzzy models that combine methods of fuzzy logic and artificial neural networks and systems, i.e. "human-like" style of considerations of fuzzy systems with training and simulation of mental phenomena of neural networks. To build a model for calculating the risk assessment of information security, it is proposed to use a fuzzy product model. Fuzzy product models (Rule-Based Fuzzy Models/Systems) this is a common type of fuzzy models used to describe, analyze and simulate complex systems and processes that are poorly formalized.

Objective. Development of the structure of a fuzzy model of quality of information security risk assessment and protection of ERP systems through the use of fuzzy neural models.

Method. To build a model for calculating the risk assessment of information security, it is proposed to use a fuzzy product model. Fuzzy product models are a common kind of fuzzy models used to describe, analyze and model complex systems and processes that are poorly formalized.

Results. Identified factors influencing risk assessment suggest the use of linguistic variables to describe them and use fuzzy variables to assess their qualities, as well as a system of qualitative assessments. The choice of parameters is substantiated and the structure of the fuzzy product model of risk assessment and the basis of the rules of fuzzy logical conclusion is developed. The use of fuzzy models for solving problems of information security risk assessment, as well as the concept and construction of ERP systems and analyzed problems of their security and vulnerabilities are considered.

Conclusions. A fuzzy model has been developed risk assessment of the ERP system. Selected a list of factors affecting the risk of information security. Methods of risk assessment of information resources and ERP-systems in general, assessment of financial losses from the implementation of threats, determination of the type of risk according to its assessment for the formation of recommendations on their processing in order to maintain the level of protection of the ERP-system are proposed. The list of linguistic variables of the model is defined. The structure of the database of fuzzy product rules – MISO-structure is chosen. The structure of the fuzzy model was built. Fuzzy variable models have been identified.

KEYWORDS: information security, fuzzy logic, risk assessment, security, ERP-system.

ABBREVIATIONS

ANFIS is an Adaptive Network-based Fuzzy Inference System;

DB is a Database;

DSTU is a State standard of Ukraine;

ERP is a Enterprise Resources Planning;

ERP-System is an Enterprise Recourses Planning System;

MISO is a Structure (Multi Inputs – Single Output);

FIS is a Fuzzy Inference System;

ARL is an acceptable risk level;

MRL is a middle risk level;

HRL is a high-risk level;

VLR is a very low risk;

LR is a low risk;

AR is an average risk;

HR is a High risk;

VHR is a Very high risk;

CVSS is a Common Vulnerability Scoring System;

NVD is a National Vulnerability Database;

CVE is a Common Vulnerabilities and Exposures.

NOMENCLATURE

R_{ij} is a Risk of the *i*-th resource in the implementation of the *j*-th threat;

A_{ij} is a Expected loss from the onetime implementation of the *j*-th threat to for the *i*-th resource;

P_j^t is a probability of occurrence of *j*-th threat;

P_{ij}^v is a Vulnerability of the *i*-th resource to the *j*-th threat;

IR is a Resource set of system;

Th is a A set of threats to the system.

A_i^V is a Value of the ist resource;

F_{ij}^e is a Impact consequences in the implementation of the *j*-th threat on the *i*-th resource, or the propensity of the *i*-th resource to the *j*-th threat;

R_i is a Risk of the *i*-th resource in the implementation of threats;

R_{ik} is a Risk of the *i*-th resource in the implementation of the *k*-th threat;

DOI 10.15588/1607-3274-2022-1-12

Th_i is a set of risks for the *i*-th resource;

R_g is a General system risk;

R_{ig} is a risk of the *i*-th resource at general system risk;

FL_i is a financial loss of the *i*-th resource;

R_i is a risk of the *i*-th resource;

Co_i is a cost of the *i*-th resource;

FL is a Total financial loss;

RL is a Risk level type;

$\min_R$ is a Minimum value of risk assessment;

$\max_R$ is a Maximum value of risk assessment;

$\Pr_1$ is a parameter, maximum value of risk assessment of acceptable type;

$\Pr_2$ is a parameter, the maximum value of the risk assessment of the average type;

$x_j (j=1,\cdots,m)$ is an Incoming Variables (can be either clear or fuzzy);

$x_j \in X_j$, X_j is an The definition area appropriate prerequisites;

y is a Fuzzy output variable;

$y \in Y, Y$ is a the definition area the conclusion;

A_{ij}, B_i is a fuzzy sets defined that are defined by X_j and Y with affiliation functions $\mu_{A_{ij}}(x_j) \in [0;1]$ and $\mu_{B_i}(y) \in [0;1]$ respectively;

p_i, q_i, r_i is a Affiliation functions options;

$k=1,\cdots,K$ is a an example from many examples of training sampling;

$x_1^{(k)}, x_2^{(k)}, \cdots, x_m^{(k)}$ are Input variable values $x_1, x_2, \cdots, x_m$;

$y^{(k)}$ is a reference value of the source variable *y* in the *k*-th example;

K is a he total number of examples, size of Training sample;

$E^{(k)}$ is a error *k*-th example from many examples of educational sample;

E is a Error;

$y^{1(k)}$ is a Installed the value of the source variable y in the *k*-th example;

ε is a installed threshold;

C is a Assessment of the criticality of information;

$C_{CC} = \max(CC_I, CC_p, CC_A, CC_O)$ are Assessment of the consequences of violations of integrity, confidentiallity, accessibility and observation for the commercial interests of the organization;

$C_{MC} = \max(MC_I, MC_p, MC_A, MC_O)$ are Assessment of the consequences of violations of integrity, confidentiality, accessibility and observation for the operational activities of the organization;

$C_R = \max(R_I, R_p, R_A, R_O)$ are Assessment of the consequences of violations of integrity, confidentiality, accessibility and observation for the organization's relationship with customers and partners.

INTRODUCTION

The basis of activity of any organization is business processes, which are determined by the goals and objectives of the entity. The business process broadly understands the structured sequence of actions to perform a certain type of activity at all stages of the life cycle of the subject of activity. Each business process has a start (login), output, and sequence of procedures that ensure that operations are grouped by the appropriate types. In general, the calculation of the risks of information security of ERP-systems should be carried out in relation to each critical business process and only on those vulnerabilities that are relevant to a particular business process, and it should be borne in mind that a number of vulnerabilities may be the same for all business processes.

Each vulnerability in the current list of vulnerabilities is correlated by a threat, the terms of which could be this vulnerability, and for each specified pair, an assessment of the probability of its occurrence and assessment of the impact of the implementation of this pair on the integrity, confidentiality, accessibility and observability is carried out.

We will use the following definitions. Probability is a conditional number that determines the likely frequency of steam threat/vulnerability. Privacy is a property of information that is that information cannot be obtained by an unauthorized user and/or process. Integrity is a property of information, which is that information cannot be modified by an unauthorized user and/or process. System integrity – system property, which is that none of its components can be eliminated, modified or added in violation of security policy. Accessibility – the property of the system resource, which is that the user and/or process, which has the appropriate powers, can use the resource in accordance with the rules established by the security policy, without waiting longer for a specified (small) period of time, that is, when it is in the form required by the user, in the place required by the user, and at the time when it is necessary. Observation – system property, which allows to record the activities of users and processes, the use of passive objects, as well as to unequivocally establish identifiers of users involved in certain events and processes in order to prevent violations of security policies and/or to ensure liability actions.

The object of the study is the development of the structure of a fuzzy model of the ERP system.

The subject of the study is neuro-fuzzy models that combine methods of fuzzy logic and artificial neural networks and systems.

DOI 10. 15588/1607-3274-2022-1-12

The purpose of the work is to improve the quality of assessment of information security risks and protection of ERP systems through the use of fuzzy neural models.

1 PROBLEM STATEMENT

Security risk assessment is an important element in the overall security risk management process, which is the process of ensuring that the organization's risk position is within acceptable limits defined by senior management and consists of four main stages: security risk assessment, testing and supervision, mitigation effects and operational security [1].

Risk managers and organizers use risk assessment to determine which risks to reduce through control and which to accept or transfer. Information security risk assessment is a process of identifying vulnerable situations, threats, the likelihood of their occurrence, the level of risks and consequences associated with organing assets, as well as control that can mitigate threats and their consequences. This process includes: assessing the likelihood of threats and vulnerabilities that are possible; calculation of the impact that can be a threat to each asset; determination of quantitative (measurable) or qualitative (described) cost of risk.

Table 1 describes the classification of technologies according to the approach used in risk assessment.

Assessment of information security risks can be divided into three stages (see Table 2): identification of risk; risk analysis; evaluation of results.

Risk assessment includes seven steps: identification of system protection facilities; identification of the threat; identification of vulnerability; control analysis; determination of probability; analysis of consequences; identification of risk.

The full risk assessment process should also include two more steps: recommendations for controlling and documenting the results.

Information risk assessment can be performed using a variety of technologies, documents or software tools. The methodology for assessing information security risks understands the systematized sequence of actions (step-by-step instructions) to be done and the tool (software product) for risk assessment at the enterprise.

Also, to assess security risks, manager documents containing theoretical descriptions can be used and provide guidelines on the risk assessment process, but no specific technologies for their implementation are provided [2–6]. At present, the following standards apply on the territory of Ukraine: ISO 27001, ISO 27002, ISO 27003, ISO 27004 and ISO 2700.

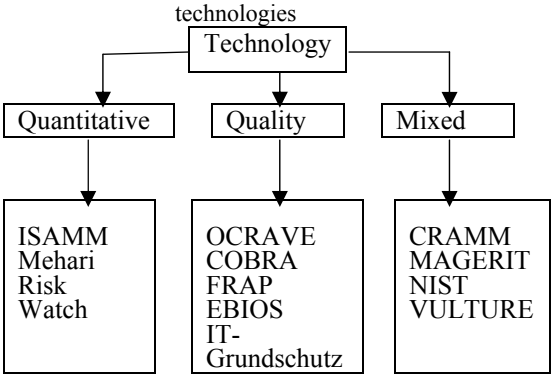
Recently, quite intensively developing methods of analysis and risk assessment, which are based on elements of fuzzy logic. Such methods allow to change the approximate table methods of rough assessment of risks to mathematical method, as well as significantly expand the possibilities of mathematical methods of risk analysis [7–11].

The mechanism of risk assessment with the help of fuzzy logic in general represents the expert system. The knowledge base of such a system complies with the rules that reflect the logic of the relationship between the input values of risk factors and the level of risk. In the simplest case, this logic is described in the table. In general, much more complex logic is used, which is designed to more accurately reflect the real relationship of factors and consequences. Such connections are formalized and described by the production rules of the "if-something" type. In addition, the mechanism of fuzzy logic involves forming levels of factor assessments and presenting them in the form of fuzzy variables. The process of forming this type of assessments in general is quite complex, because it requires a large number of sources of information, taking into account their quality and use of expert experience.

2 REVIEW OF THE LITERATURE

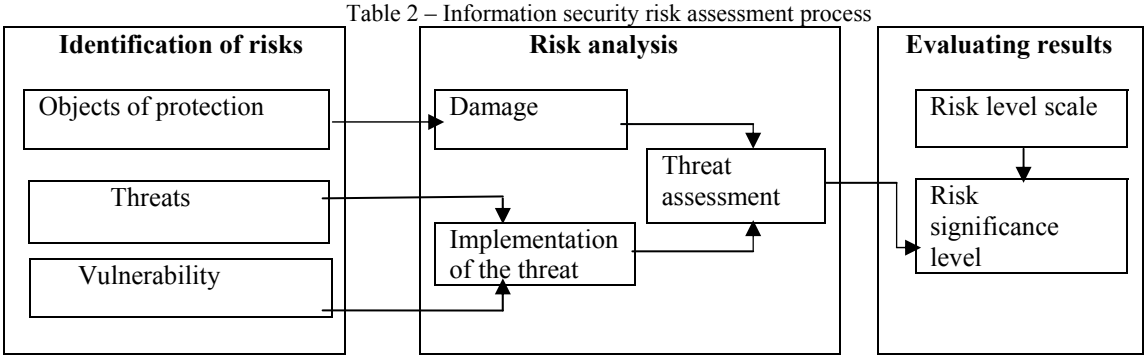
The security risk analysis study begins in the mid-1980s, and in the early 90s R. Baskerville identified risk analysis checklists for tools used to design information system security measures [11]. Over time, complex tools are developed to analyze risks, such as: Facilitated Risk Assessment Process [12]; The Operationally Critical Threat, Asset, and Vulnerability Evaluation) [13]; CORAS [14]; Is Risk Analysis Based on Business Model [15]; Information Security Risk Analysis Method [16]; Risk Watch method [17]; Consultative Objective and Bifunctional Risk Analysis [18]; CRAMM [19].

Table 1 – Information security risk assessment technologies



DOI 10. 15588/1607-3274-2022-1-12

Table 2 – Information security risk assessment process



Also, since the early 2000s, some other methods of modeling security risks, which have provided good indicators and have been commonly titled "soft computing models", including the grey relactional approach, have also been used in the research industry, Fuzzy number arihmetic, Information entropy, Fuzzy weighted average aproach, Fuzzy measure and Evidence theory, fuzzy Analysis of Hierarchy Process method.

The development and application of soft computing and hybrid models are considered to be modern areas of research to assess information security risks.

Soft computing components include: Neural networks – computational systems that assess the risks of information security through similar functioning of biological neural networks and learning tasks (gradually improving their performance of these networks), considering examples, in general, without special programming for the task; Rough sets – an effective mathematical analysis tool to address uncertainty in the field of solution analysis; Grey sets; Fuzzy systems – based on the algorithm for obtaining fuzzy conclusions based on fuzzy preconditions; Generic algorithms belong to the largest class Evolutionary algorithms and generate solutions to optimization problems using methods borrowed from the theory of evolution, such as inheritance, mutation, selection and crossover; Support vector machine – the data analysis method for classification and regression analysis using managed learning models is used when input is either not defined or when only some data is determined by their preprocessing; Bayesian network – used to identify cause and effect relationships of risk factors and predict the likelihood of security risk.

Hybrid models represent a combination of two or more technologies to develop robust risk assessment and information systems. The most common hybrid model is the neuro-fuzzy network.

To determine the level of risk, it is advisable to use the apparatus of the theory of fuzzy sets, which allows you to describe vague concepts and knowledge, operate them and draw vague conclusions. The theory of fuzzy sets is used precisely to solve problems in which inputs are unreliable and poorly formalized, as in the case of the problem solved in this work. To assess the risk, it is appropriate to use the mechanism of a vague logical conclusion – obtaining a conclusion in the form of a fuzzy set corrtsponding to the current values of input variables, using a fuzzy knowledge base and fuzzy operations.

There are developed models of fuzzy conclusion of Mamdani, Sugeno, Larsen, Tsukamoto [20]. Most often, Mamdani and Sugeno algorithms are used in practice. The main difference between them is the way to set the values of the source variable in the rules that constitute the knowledge base. In systems like Mamdani, the values of input variables are set by fuzzy terms, in systems like Sugeno – as a linear combination of input variables. For tasks in which identification is more important, it is ad.

3 MATERIALS AND METHODS

To build a structure a model for calculating information security risk assessment, it is proposed to use Rule-Based Fuzzy Models/Systems.

Under the Rule-Based Fuzzy Models/Systems understand the agreed a lot of individual fuzzy product rules of the type "if A, then B" where A is the prerequisite (parcel, antecendent) of a certain rule, and B – the conclusion (action, consecvent) of the rule in the form of fuzzy statements. The model is designed to determine the degree of truthfulness of the conclusions of fuzzy product rules. The degree of truth is determined on the basis of preconditions with a certain degree of truthfulness of the relevant rules.

When building a fuzzy product model, the following components are determined: method of fuzzy withdrawal of conclusions; database of fuzzy product rules; fuzzyfication input procedure; procedure aggregation of the degree of truthfulness of preconditions for each of the fuzzy product rules; activation procedure for each of the fuzzy product rules; the procedure of liquidation of activated conclusions of all fuzzy product rules according to each output variable; defuzzyfication procedure to clarity on each consiluled output variable; the procedure for parameters optimization of the final base of fuzzy rules.

At present, many different types of fuzzy product models are offered on the basis of different combinations of these components.

Rule-Based Fuzzy Models/Systems are used in solving a number of problems in which information about the system, its parameters, as well as the inputs, outputs and states of the system is unreliable and poorly formalized. Together with the advantages of describing the model in a language close to natural, in the versatility and efficiency of the model, Rule-Based Fuzzy Models / Systems are characterized by certain disadvantages: the wording of the original set of fuzzy rules is carried out with

DOI 10. 15588/1607-3274-2022-1-12

the help of an expert, so it may be incomplete or contradictory; the choice of the type and parameters of the functions of belonging in fuzzy statements of the rules is subjective; automatic acquisition of knowledge cannot be performed.

To eliminate these shortcomings, it is proposed to use an adaptive fuzzy production model, which in the process and on the results of functioning corrects both the composition of the rules in the base and the parameters of the functions of belonging, as well as to implement various components of this model on the basis of neuronet technology.

Determine the incoming and outgoing parameters of the model.

To build a risk assessment calculation model, we will use the risk factor ratio according to the formulas (1, 2) [10].

$$R_{ij} = A_{ij} \cdot P_j^t \cdot P_{ij}^v,\ i \in IR,\ j \in Th. \quad (1)$$

Under the expected damage from a one-time implementation of the threat we understand the cost (or value) of the asset, which is mathematically expressed as follows:

$$A_{ij} = A_i^V \cdot F_{ij}^e, i \in IR,\ j \in Th. \quad (2)$$

Taking into account (1) and (2), we obtain the general ratio of factors for risk assessment:

$$R_{ij} = A_i^V \cdot F_{ij}^e \cdot P_j^t \cdot P_{ij}^V, i \in IR,\ j \in Th. \quad (3)$$

Since many risks can be identified for each information resource (one to all), the assessment of the total risk by the information resource will be defined as the maximum risk assessment of the resource:

$$R_i = \max(R_{ik}),\ k \in Th_i. \quad (4)$$

In turn, the assessment of system risk will be defineed as the maximum assessment among resource risk assessments:

$$R = \max(R_i),\ i \in IR. \quad (5)$$

The amount of financial damage for the information resource will be determine as the product of the risk of the information resource on the cost of the resource:

$$FL_i = R_i \cdot Co_i,\ i \in IR. \quad (6)$$

In turn, the total financial loss will be determined as the amount of financial losses on all resources:

$$FL = \sum_i FL_i,\ i \in IR. \quad (7)$$

We will apply a linguistic approach to the description of information security risk factors. Suppose as the values of factors and characteristics of relations between them not only quantitative assessment, but also qualitative, sentences of natural language. Then this approach will provide a quantitative description of the elements of the model in the conditions of vague information about the value of the risk level, the cost of the resource, the impact of the consequence of, the likelihood of a threat, the vulnerability of resource protection and ways to avoid negative impact from the implementation of risks.

Each risk factor of information security and the risk itself will be described by linguistic variables $X \in \overline{X}$, Wheel the set of linguistic variables of the model $\overline{X}$ is: $\overline{X}$ = {“Resource Price”, “Impact of the consequence”, “Probability the emergence of Threat”, “Resource Vulnerability”, “Risk”}.

The list of linguistic variables of the model corresponding to the risk factors is shown in Table 3.

Thus, information security risk assessment can be expressed as:

$$Y = f_Y(X_1, X_2, X_3, X_4).$$

Based on the analysis [21] and the formed ratio of risk factors (3) for the assessment of each of the risks, a fuzzy model with four input parameters (X_1, X_2, X_3, X_4) and one Y output (MISO structure [22]) is proposed. The number of input parameters is selected according to the number of factors influencing the degree of risk (3). Table 3 shows the structure of the system of fuzzy conclusions for the selected model.

Table 3 – The list of linguistic variables of the model

List	Name of linguistic variable
X_1	Resource Price
X_2	Impact of the consequence
X_3	Probability the emergence of Threat
X_4	Resource Vulnerability
Y	Risk

To maintain the level of security of the ERP system, it is necessary to determine what risks, according to the level of their assessment – risk level (RL), require processing according to certain recommendations. To do this, we will introduce 3 types of risk levels:

– acceptable risk – ARL – will be considered insignificant, the processing of such a risk is not required;

– medium risk – MRL – recommended for processing in order to minimize it;

– high risk-HRL – we will consider it essential and its processing is mandatory.

Determination of the type of risk will be carried out as follows:

DOI 10. 15588/1607-3274-2022-1-12

$$RL=\begin{cases} ARL, R_{ij}\in(\min_R;\Pr_1); \\ MRL, R_{ig}\in(\Pr_1;\Pr_2); \quad i\in IR, j\in Th, \\ HRL, R_{ij}\in(\Pr_2;\max_R). \end{cases} \tag{8}$$

Parameters – the maximum value of the assessment of acceptable and medium risk – $[\Pr_1]$ and $[\Pr_2]$ respectively – are set by experts.

The scheme for processing the result of risk assessment is shown in Table 4.

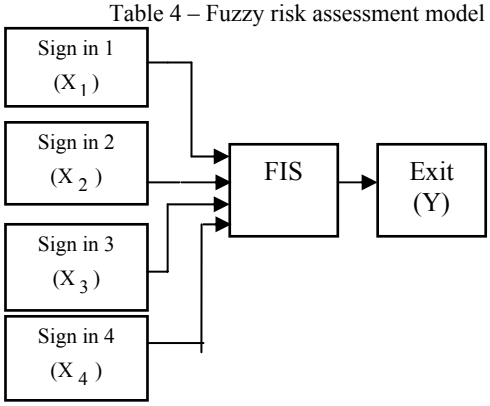
We will create a structure and build bases of fuzzy product rules.

The structure of the rules should correspond to the structure of the model, namely the number of fuzzy statements in the prerequisites and conclusions. The database of rules that has the structure of MISO, in general, has the following rule structure [22].

$$\text{P}_i: \text{If } x_1 \text{ is } A_{i1} \text{ and } \ldots \text{ and } x_j \text{ is } A_{ij} \text{ and } \ldots \text{ and } x_m \text{ is } A_{im}, \text{ then } y \text{ is } B_i. \tag{9}$$

When creating a fuzzy product model, both a priori data coming from experts and data obtained as result of measurements can be used.

Table 4 – Fuzzy risk assessment model



In the first case, if there is no need to agree on the opinions of experts, it is assumed that the tasks of ensuring completeness and inconsistency of the database of fuzzy rules are solved in advance. If only experimental data are known, these tasks can be attributed to the tasks of system identification. In practice, there may also be a mixed case when the initial database of fuzzy rules is built on the basis of heuristic assumptions, and its clarification is carried out using experimental data.

ANFIS, the adaptive network fuzzy output system proposed by Chang in 1992, will be used to represent the fuzzy production model and algorithm of fuzzy output in the form of a fuzzy product network [23].

Since the fuzzy ANFIS product network is presented as multilayer structure with a direct signal propagation, and the value of the source variable can be changed by adjusting the parameters of layer elements, then to teach this network you can use an algorithm for reverse spread ing the error, which belongs to the class of classic gradient algorithms.

Consider the problem of fuzzy neural production network of anfis type, which implements the algorithm of fuzzy output of Takagi-Sugeno [24] (see Figure 1).

Let the rules of this form be set:

P1: If x_1 is A_{11} and x_2 is A_{12} then $y_1 = a_1 x_1 + b_1 x_2$;

P2: If x_1 is A_{21} and x_2 is A_{22} then

$$y_2 = a_2 x_1 + b_2 x_2. \tag{10}$$

The structure of the fuzzy neural production network of ANFIS type, which implements the algorithm of fuzzy output of Takagi-Sugeno (according to the example) is shown in Fig. 2 [23].

Layer 1. The outputs of the elements of this layer are $\mu_{A_{ij}}(x_j)$ the values of the functions of the affiliation at specific (specified) values of input variables. For example, circular functions hare the form of:

$$\mu_{A_{ij}}(x_j) = \exp\left[-\frac{1}{2}\left(\frac{x_j - a_{ij}}{b_{ij}}\right)^2\right]. \tag{11}$$

Layer 2. Elements of the second layer perform aggregation of the truth levels of the prerequisites of each base rule in accordance with the T-norm operation, which uses the operation minimum (4) [20] according to the rules:

$$\begin{aligned} a_1 &= \min\{A_{11}(x_1), A_{12}(x_2)\}, \\ a_2 &= \min\{A_{21}(x_1), A_{22}(x_2)\}. \end{aligned} \tag{12}$$

Layer 3. Elements of this layer normalize and lead these results to a type convenient for calculating the output of a fuzzy network. Calculation β_i -normalized values α_i are performed as follows:

$$\beta_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}, \beta_2 = \frac{\alpha_2}{\alpha_1+\alpha_2}. \tag{13}$$

Layer 4. Elements in this layer calculate function values:

$$\begin{aligned} y_1^{/} &= (p_1 x_1 + q_1 x_2 + r_1), \\ y_2^{/} &= (p_2 x_1 + q_2 x_2 + r_2). \end{aligned} \tag{14}$$

Layer 5. Elements of this layer allow you to form a defaziated value at the output of the network, which is formed as follows:

$$y^{/} = \beta_1 y_1^{/} + \beta_2 y_2^{/}. \tag{15}$$

DOI 10. 15588/1607-3274-2022-1-12

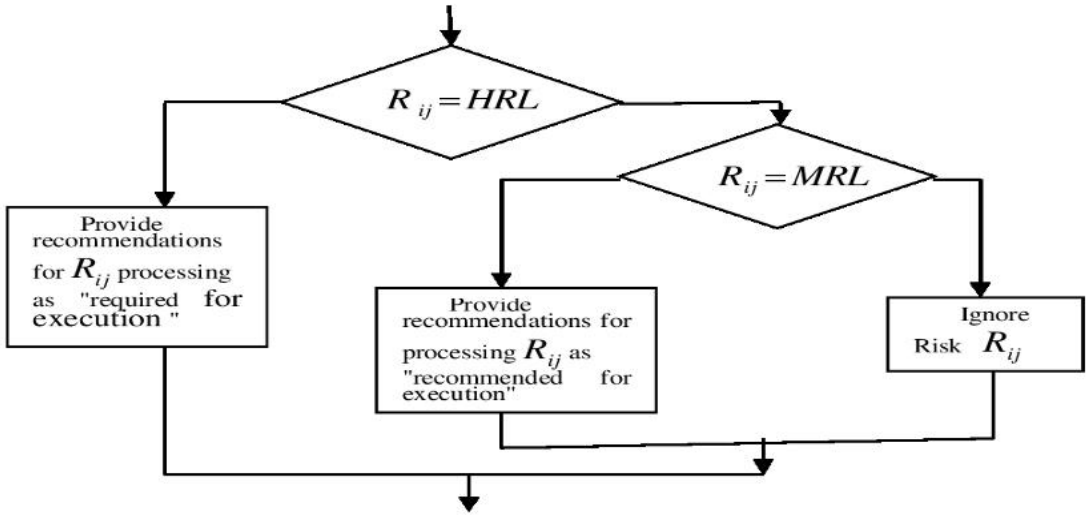


Figure 1 – Scheme for processing the result of risk assessment

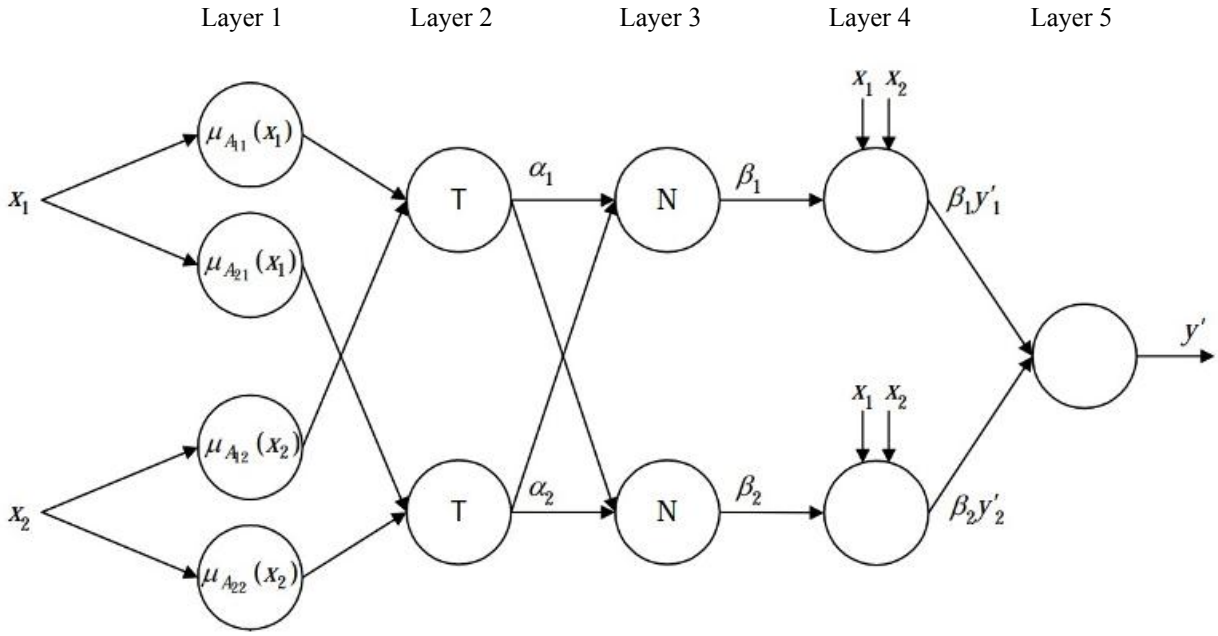


Figure 2 – The structure of the fuzzy ANFIS neural production network, which implements the Sugeno fuzzy output algorithm

Parametric layers fuzzy ANFIS neural production network, that is, the layers, the parameters of the elements in which will be adjusted during the learning process, are the first and fourth, and the parameters configured in the learning process are:

– in the first layer – nonlinear parameters of the affiliation functions $\mu_{A_{ij}}(x_j)$ fuzzy sets of preconditions of the rules;

– in the fourth layer – nonlinear parameters p_i, q_i, r_i affiliation functions $\mu_{B_i}(y)$ fuzzy sets of rule conclusions.

The Picks for learning the network consists of many examples and has the form of:

$$(x_1^{(k)}, x_2^{(k)}, \cdots, x_m^{(k)}, y^{(k)}). \qquad (16)$$

Since the fuzzy ANFIS product network is presented as multilayer structure with a direct signal propagation, and the value of the source variable can be changed by adjusting the parameters of layer elements, then to study this network we use an algorithm for reverse spreading the error, which belongs to the class of gradient algorithms.

Network training continues (iteratively repeats the procedure for adjusting the values of all parameters) as long as [21]:

– or the error function value for each sample example does not exceed some set threshold:

$$E^{(k)} < \varepsilon, k = 1, \cdots, K. \qquad (17)$$

DOI 10. 15588/1607-3274-2022-1-12

– or assessment of the average total error of a fuzzy product model, taking into account all examples of the educational sample does not exceed some established threshold:

$$E = \frac{1}{k}\sum_{k=1}^{K}(y^{/(k)} - y^{(k)})^2 < \varepsilon. \quad (18)$$

4 EXPERIMENTS

Let's define a linguistic variable Y "Risk". To evaluate the linguistic variable Y, we will use the term set T(Y) of five quality thermals: $T(Y)$={"Very level risk (VLR); "Low risk (LR) "; "Average risk (AR)"; "High risk HR)"; "Very high level of risk (VHR)"}.

Definition Area of E_Y of the linguistic variable Y will be set at the interval [0,100]. Table 5 shows the scale for assessing the level of risk.

Taking into account the selected area of determining the risk assessment of information security when determining the type of risk to make recommendations for its reducti-on according to the formula (6), we will use the following values:

Table 5 – Y Risk Assessment Scale

Risk assessment	Risk level
0–20	VLR
20–40	LR
40–60	MR
60–80	HR
80–100	VHR

$\min_R = 0, \max_R = 100.$

Consider the definition identifying threats and assessing the likelihood of threats. The main security threats of ERP systems include deliberate actions of violators, for example, criminals, spies, saboteurs, or offended persons from among the personnel of the organization [25].

1. According to the results of the actions of violators:

– threat of information leakage;

– threat of information modification; – threat of loss of information.

2. Based on the motives of the violators: unintentional; deliberate.

According to the Normative Document in the Field of Technical Information Protection (GNI TZI) 2.5-004-99 [25] in the risk assessment model we will consider threats of the following four types in accordance with the properties of information security:

– threats related to unauthorized acquaintance with information and pose threats to the confidentiality of information;

– threats related to unaut-horized modification of information and pose threats to the integrity of information;

– threats related to violation of the possibility of using the system or information that is processed and poses threats of violation of the availability of information;

– threats related to the violation of the possibility of surveillance, managing and controlling user activity, the possibility of legality of access, capabilities and capabilities to perform the functions of a complex of means of protection and pose a threat of violation of the observation of information.

When analyzing the negative consequences of influencing the ERP system of different types of information th-reats, as a rule, their following categories are considered [26].

Refusals and hardware failures and/or network failures, emergencies and other events occurring without the participation of personnel; unintentional or erroneous actions of administrators, users, system operators or other types of personnel; unauthorized access by violators to the information that is generated, processed and stored in the ERP system, for example, information that:

– perform management and decision making – information of users of the ERP system; provides equipment management of ERP-system; allows you to implement business processes and technologies of information processing in the ERP system.

The "subjective" and "objective" probability of a threat is calculated by expert methods using mathematical methods. The frequency of threats can be determined by quantitative indicator in accordance with the number of cases of threat per year.

To evaluate the linguistic variable X_3 "Threat probability level", we will use the term set $T(X_3)$ of five quality therms: $T(X_3)$={Very low probability of threat (VLT); Low probability of threat (LT); Average threat probability (MT); High probability of threat (HT); Very high probability (VHT).

Definition Area E_{X_3} of the linguistic variable X_3 beset at the interval [0, 05; 365].

Table 6 provides a scale for assessing the level of threat probability in accordance with the frequency of threats per year.

When evaluating the linguistic variable X_4 "Resources Vulnerability", we will rely on the common vulnerability assessment system (CVSS), which makes it possible to fix the basic characteristics of the vulnerability and create a numerical score that reflects its criticality [27]. CVSS is a free and open industry standard for assessing the severity of a computer system security vulnerability, allowing users to prioritive resources according to threat. The CVSS assessment system consists of three metrics [26]: basic metric – reflects the basic qualities and characteristics of the vulnerability; temporary metrics – reflects the following characteristics of the vulnerability, which change over time, develop during a vulnerable period; contextual metric – displays the characteristics of the vulnerability that are unique to the user environment.

DOI 10. 15588/1607-3274-2022-1-12

Each metric group has a specific numerical score (rating) in the range from 0 to 10 and a period representing the value of all metrics in the form of a block of text.

To obtain highquality vulnerability metrics, we will use the National Vulnerability Database (NVD) assessment system. NVD is an information database of the U.S. National Standardization Authority, the U.S. Government supported National Institute of Standards and Technology, that collaborates with the Common Vulnerabilities and Exposures (CVE) database, which represents a dictionary of commonly used names (such as CVE identifiers) for publicly available information security vulnera-bilities. In the NVD database, the security level values of the vulnerability are calculated by values from 0 to 10 (according to CVSS) and are described linguistically by the term None, Low, Medium, High and Critical.

According to the linguistic therms of the NVD database, we will use the $T(X_4)$ term set of four quality therms to evaluate the linguistic variable X_4 "Resource Vulnerability":

$T(X_4)$ = {Low vulnerability (LV); Medium vunerability (MV); High vulnerabilidad (HV); Critical vulnerability (CV).

Definition Area E_{X_4} of the linguistic variable X_4 set at the interval [0,10].

Table 7 describes NVD vulnerability scores by points and linguistically, description of the impact of exploitation, and corresponding levels of resource vulnerability according to the term sets $T(X_4)$.

We will determine the consequences of violation of the integrity, confidentiality, accessibility and observation of information in such important areas of activity of the organization as: – commercial concernment (CC); – management control (MC); relation with clients and partners – relations (R).

The results of the assessment of the consequences of Violation of integrity, confidentiality, accessibility and observation of information in the spheres of activity of the organization are given in Table 8.

Table 6 – Threat probability level assessment scale (X_3)

Frequency	Probability of occurrence a threat for a certain period	Level
0,05	threat is almost never realized	VLT
0,6	approximately 2–3 times in five years	VLT
1	approximately once a year and less (180 <in> 366 (days))	LT
2	approximately 1 time in six months (90 <in > 180 (days))	LT
4	approximately 1 time in 3 months (60 <in > 90 (days))	MT
6	approximately 1 time in 2 months (30 <in > 60 (days))	MT
12	approximately 1 time per month (15 <in > 30 (days))	HT
24	approximately 2 times a month (7 <In > 15 (days))	HT
52	approximately 1 time per week (1 <In > 7 (days))	VHT
365	Daily (1 <In > 7 (Hours))	VHT

Table 7 – Resource Vulnerability Rating Scale

Level by NVD	Score by NVD	Description of the vulnerability level	Vulnerability level
None	**0.0**	Vulnerability has no effect on resource	
Lov	**0.1–3.9**	A vulnerability that has little impact on the resource does not Affect the availability, integrity and confidentiality of information	**LV**
Medium	**4.0–6.9**	A vulnerability that may have some impact on the resource but has a complexity of implementation or does not cause serious consequences. It is possible to access confidential information, change some information, but there is no control over the information, or the scale of losses is small. Resource availability failures occur	**MV**
High	**7.0–8.9**	A vulnerability that has a significant impact on the resource, possible access to confidential information, changes in information and control over information. Significant resource availability failures and performance reductions	**HV**
Critical	**9.0–10.0**	Vulnerability, the consequence of the exploitation of which has a serious impact on the resource: complete loss of availability and integrity of information, full disclosure of confidential information	**CV**

Table 8 – Assessment of the consequences of violation of information properties in the spheres of activity

	Spheres of activity of the organization		
Information Resource Property	**Commercial concernment (CC)**	**Management control (MC)**	**Relationships with clients and partners (R)**
Integrity	CC_I	MC_I	R_I
Confidentiality	CC_I	MC_P	R_P
Accessibility	CC_A	MC_A	R_A
Observability	CC_O	MC_O	R_O

DOI 10. 15588/1607-3274-2022-1-12

To assess the consequences of the threat, we will use a quantitative assessment of the impact on certain properties of information (integrity, confidentiality, accessibility and observation), as proposed by the NBU Methodological Recommendations (National Bank of Ukraine).

The values of assessments of the consequences of violation of integrity, confidentiality, availability and observation of information for commercial interests (CC), management control (MC) and customer-to-partner relationships (R) will be within the range of integer values [1,5].

We will calculate the impact assessment (CA) for each property of the information.

Assessment of the consequences of integrity violation:

$$CA_I = \max(CC_I, MC_I, R_I).$$

Assessment of the consequences of a privacy violation:

$$CA_P = \max(CC_P, MC_P, R_P).$$

Assessment of the consequences of accessibility violations:

$$CA_A = \max(CC_A, MC_A, R_A).$$

Assessment of the consequences of observational violations:

$$CA_O = \max(CC_O, MC_O, R_O).$$

The implementation of the threat can affect several properties at once, so it is necessary to determine thee general assessment of the consequences of violation of the properties of information:

$$CA = \max(CA_I, CA_P, CA_A, CA_O). \quad (19)$$

To evaluate the linguistic variable X_2 "Impact the consequence of", will use the term set $T(X_2)$ of five quality therms:

$T(X_2)$= {Very low consequences (VLC); Low consequences (LC); Medium consequences (MC); Significant consequences (SC); Very big consequences (VBC)}.

The impact level assessment scale is shown in Table 9.

Table 9 – Impact Level Assessment Scale Consequences

Score	Impact Level Description	Level of impact
1	Very low consequences	VLC
2	Low consequences	LC
3	Medium consequences	MC
4	Significant consequences	SC
5	Very big consequences	VBC

The value of information will be defined as the relationship between the type of confidentiality and criticality – criticality (C) of the information. Value estimation is formed as the sum of points corresponding to each type and level of criticality of information. Estimates of the value of information are given in Table 10.

The criticality of the information will be determined, taking into account the assessment of the consequences of violation of the properties of information (see Table 2) by the formula

$$C = C_{CC} + C_{MC} + C_R. \quad (20)$$

To evaluate the linguistic variable X_1 "Resource price", we will use the term set $T(X_1)$ of three highquality therms:

Basis of the development of information risk management systems.

$T(X_1)$= {Low Price (LP); Average Price (AP); High Price (HP)}.

The Definition Area of E_{X_1} of the linguistic variable X_1 be set at the interval [19]. The scale for assessing the value level of information is presented in Table 11.

Table 11 – Information Value Assessment Scale

Price	Description of Price level	Level of Price
4	Low Price	LP
11	Average Price	AP
19	High Price	HP

Table 10 – Definition of value assessment of information

	Criticality of information (C)		
Type of information	Insignificant (1–3 points)	Significant (4–9 points)	Critical (10–15 points)
Open (1 point)	2–4	5–10	11–16
For internal use (2 points)	3–5	6–11	12–17
Confidential (3 points)	4–6	7–12	13–18
Strictly Confidential (4 points)	5–7	8–13	14–19

DOI 10. 15588/1607-3274-2022-1-12

5 RESULTS

This article developed a fuzzy risk assessment model of the ERP system and performed the following stages of Development: a list of factors influencing information security risk is selected; suggested methods for assessing the risk of information resources and ERP-systems in general, assessing financial losses from the implementation of threats, determining the type of risk according to its assessment to form recommendations for their processing in order to maintain the level of security of the ERP-system; the list of linguistic variables of the model is determined; the structure of the base of fuzzy product rules – MISO-structure was chosen; the structure of the fuzzy model was built; fuzzy model variables are defined; the principles of construction of systems of fuzzy logical conclusion and neuro-fuzzy models, use of fuzzy models to solve problems of risk assessment of information security are considered. The concept, principles of construction, functioning and requirements for information security of ERP systems are considered, problems of their safety and vulnerability are analyzed.

According to the results of the review, the main factors influencing the risk assessment are determined, the choice of parameters of a fuzzy product model for risk assessment and the structure of the rules base of a fuzzy logical conclusion is substantiated. Adaptive neuro-fuzzy product model of risk assessment of information security threats is developed.

It is proposed to use a linguistic approach to describe the main factors influencing the assessment of risks, variables and fuzzy variables to assess their qualities, as well as a system of qualitative assessments. The choice of parameters was substantiated and the structure of a fuzzy product model for risk assessment and the basis of the rules of a fuzzy logical conclusion were developed.

As a result, the developed adaptive neuro-fuzzy product model for risk assessment of information security of ERP systems allows to perform risk assessment on four factors: resource value, impact of impact on resource, probability of threat and vulnerability of the resource.

The obtained risk assessments can be used both to assess the risks of information security of ERP-system resources and to the general risk of information security of the ERP system.

The use of a linguistic approach ensures the possibility of using quantitative description of both all and individual elements of the model, provided that there is only infmation about the value of fuzzy information security risk factors, which provides opportunities, if necessary, to separate and rank risk factors and their consequences. Such actions may be useful in determining ways to avoid and /or reduce the negative impact of risk.

The use of neuro-fuzzy system components gives the model flexibility. Setting up the model by training in accordance with the obtained knowledge base allows you to perform risk reassessment in case of changes in the values of factors, changes in the product base of rules or the emergence of new risks. This provides an opportunity to shape and adapt the model to a specific ERP system.

6 DISCUSSIONS

Violation of information security, including noncompliance with regulatory standards, can lead to financial and reputational consequences that are best avoided for any organization, regardless of size, scope or form of ownership.

The operating procedures and business applications that support them must be strategically managed and monitored to ensure the integrity, availability and confidentiality of the data that the organization owns.

Currently, the vast majority of organizations rely on ERP-Systems to implement business processes and integrate financial data. The ERP system is an application system that implements a strategy of comprehensive resource planning that integrates the company's business processes and financial data into one platform. Integration provides better quality and availability of information, but it also increases the risk of fraud from within the organization by users and malicious attacks from outside. This dependency increases the security value of the ERP system to protect your organization's information assets.

A key aspect of any security strategy is the ability to achieve a level of security that adequately demonstrates the organization's commitment to information security and data security regulations collected from its customers and partners. Too little security increases the risk of violations, while too much can lead to unnecessary costs for information technology, software and hardware, deteriorating system performance, and slowing down business processes. There is no optimal security solution for any ERP-system. Each organization needs to assess risks and set goals related to their environment and the type of information it processes.

The peculiarity of risk assessment tasks is that most of the data on risk factors has signs of imperfection and uncertainty: contradiction, inaccuracy, unreliability or incompleteness, are nonlinear and dynamically variable. For effective assessment in case of uncertainty of input data, fuzzy logic methods and neuro-fuzzy networks are used to use linguistic variables and statements to describe risk factors and be adaptive at the expense of the neuro-network component.

ACKNOWLEDGEMENTS

The work was performed at the Department of Information and cybernetic security of the State University of telecommunications within scientific researches conducted by the department.

DOI 10. 15588/1607-3274-2022-1-12

REFERENCES

1. Leighton J. Security Controls Evaluation, Testing and Assessment Handbook. Syngress, 2016, 678 p.
2. Rescher N. «Many-Valued Logic», Mc.Graw-Hill. New York, 1969. DOI:10.2307/2272880
3. Rosser J. B., Turquette A. R. Many-Valued Logics, North Holland. Amsterdam, 1952.
4. Common Vulnerability Scoring System version 3.1: Specification Document. CVSS Version 3.1 Release [Elektronnyi resurs], *Forum of Incident Response and Security Teams*. Rezhim dostupu: https://www.first.org/cvss/ specification-document.
5. Abhishek kumar srivastav, Irman Ali, Shani Fatema. A Quantitative Measurement Methodology for calculating Risk related to Information Security, *IOSR Journal of Computer Engineering (IOSR-JCE),* Volume 16, Issue 1, Ver. IX (Feb. 2014), pp. 17–20.
6. Hayashi Y., Imura A. Fuzzy neural expert system with automated extraction of fuzzy If-Then rules from a trained neural network, *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*, 1990, pp. 489–494. DOI.1109/ISUMA.1990.151303
7. Buckleya J. J., Hayashi Y. Fuzzy neural networks: a survey, *Fuzzy sets and systems,* 1994, Vol. 66, Issue 1, pp. 1–13. https://doi.org/10.1016/0165-0114(94)90297-6
8. Hendrawirawan D. Tanriverdi H., Zetterlund C. ERP Security and Segregation of Duties Audit: A Framework for Building an Automated Solution, *Information systems control journal*, 2007, Vol. 2, 4 p. ISACA. All rights reserved. www.isaca.org
9. Nieto-Morote A., Ruz-Vila A. F. Fuzzy approach to construction Project risk assessment, *International Journal of Project Management*, 2011, Vol. 29, Issue 2, pp. 220–231.
10. Kozhukhivskyi A. D., Kozhukhivska O. A. ERP-System Risk Assessment Methods and Models (Tekst), *Radio Electronics, Computer Science, Control*, 2020, No. 4(55), pp. 151–162. DOI 10.15588/1607-3274-2020-4-15.
11. Baskerville R. An analysis survey of information system security design methods: Implications for Information Systems Development, *ACM Computing Survey*, 1993, pp. 375–414.
12. Peltier T. R. Facilitated risk analysis process (FRAP), *Auerbach Publication,* CRC Press LLC, 2000, 21 p.
13. Alberts C., Dorofee A. Managing Information Security Risks: The Octave Approach. Addison-Wesley Professional, 2002, 512 p.
14. Stolen K., den Braber F., Dirmitrakos T. Model-based risk assessment – the CORAS approach [Elektronnyi resurs], 2002. Rezhim dostupu, http://folk.uio.no/nik/2002/Stolen.pdf
15. Suh B., Han I. The IS risk analysis based on business model, *Information and Management*, 2003, Vol. 41, No. 2, pp. 149–158.
16. Karabacak B., Songukpinar I. ISRAM: Information security risk analysis method, *Computer & Security, March,* 2005, pp. 147–169.
17. Goel S., Chen V. Information security risk analysis – a matrix-based approach [Elektronnyi resurs], *University at Albany*, SUNY, 2005. Rezhim dostupu: https://www.albany.edu /~goel/publications/goelchen2005.pdf
18. Elky S. An introduction to information system risk management [Elektronnyi resurs], *SANS Institute InfoSec Reading Room*, 2006. Rezhim dostupu: https://www.sans.org/reading-room/whitepapers/auditing/introduction-information-system-risk-management-1204.
19. Yazar Z. A. A Qualitative risk analysis and managment tool – CRAMM [Elektronnyi resurs], *SANS Institute InfoSec Reading Room*, 2011, Rezhim dostupu: https://www.sans.org/reading-room/whitepapers/auditing/qualitative-risk-analysis-management-tool-cramm-83
20. Korchenko A. G. Building information protection systems on fuzzy sets. Theory and practical solutions. Kyev, MK-Press, 2006, 320 p.: IL.
21. Security issues in ERP. Security, Audit and Control Features SAP ERP 4th Edition, Audit Program. Isaca, 2015, 574 p.
22. A Complete Guide to the Common Vulnerability Scoring System. Forum of Incident Response and Security Teams (June2007). Rezhim dostupu: http://www.first.org/cvss/cvss-guide.pdf
23. Polyakov A. ERP Security Deserves Our Attention Now More Than Ever [Elektronnyi resurs], Forbes, 2017. Rezhim dostupu: https://www.forbes.com/sites/forbestechcouncil/2017/07/07/erp-security-deserves-our-attention-now-more-than-ever/.
24. "NVD Common Vulnerability Scoring System Support v2". National Vulnerability Database. National Institute of Standards and Technology. Retrieved March 2, 2013.
25. Jang J.-S. R. ANFIS: Adaptive Network – based Fuzzy Inference System, *IEEE Trans. On Syst. Man and Cybernetics*, 1993, Vol. 23, No. 3, pp. 665– 685.
26. National vulnerability database Release [Elektronnyi resurs], *National Institute of Standards and Technology.* Rezhim dostupu: https://nvd.nist.gov
27. National vulnerability database Release. Vulnerability Metrics [Elektronnyi resurs], *National Institute of Standards and Technology.* Rezhim dostupu: https://nvd.nist.gov/vuln-metrics/cvss

Received 08.11.2021.
Accepted 16.12.2021.

УДК 004.94

РОЗРОБКА НЕЧІТКОЇ МОДЕЛІ ОЦІНКИ РИЗИКІВ ДЛЯ ERP-СИСТЕМИ

Кожухівський А. Д. – д-р техн. наук, професор, професор кафедри інформаційної та кібернетичної безпеки Державного університету телекомунікацій, Київ, Україна.

Кожухівська О. А. – д-р техн. наук, доцент кафедри інформаційної та кібернетичної безпеки Державного університету телекомунікацій, Київ, Україна.

АНОТАЦІЯ

Актуальність. Оскільки оцінка ризиків інформаційної безпеки є складним і повним процесом невизначеності, а невизначеність є основним фактором, що впливає на ефективність оцінки, доцільно використовувати нечіткі методи та моделі, які є адаптивними до необчислюваних даних. Формування розпливчастих оцінок факторів ризику є суб'єктивним, а оцінка

DOI 10. 15588/1607-3274-2022-1-12

ризиків залежить від практичних результатів, отриманих у процесі обробки ризиків загроз, які вже виникли під час функціонування організації та досвіду фахівців з інформаційної безпеки. Тому доцільним буде використання моделей, що здатні адекватно оцінювати нечіткі фактори та мають можливість корегування їх впливу на оцінку ризику. Найбільші показники ефективності для вирішення таких задач мають нейро-нечіткі моделі, що комбінують методи нечіткої логіки та штучних нейронних мереж і систем, тобто «людиноподібного» стилю міркувань нечітких систем з навчанням та моделюванням розумових явищ нейронних мереж. Для побудови моделі розрахунку оцінки ризику інформаційної безпеки пропонується використовувати нечітку продукційну модель. Нечіткі продукційні моделі (нечіткі моделі/системи на основі правил) це поширений тип нечітких моделей, які використовуються для опису, аналізу та моделювання складних систем і процесів, що слабо формалізуються.

Мета роботи – розробка структури нечіткої моделі оцінки ризиків інформаційної безпеки та захисту систем ERP шляхом використання нечітких нейронних моделей.

Метод. Для побудови структури моделі розрахунку оцінки ризику інформаційної безпеки пропонується використовувати нечітку продукційну модель. Нечіткі продукційні моделі це загальний вид нечітких моделей, які використовуються для опису, аналізу та моделювання складних систем і процесів, що слабо формалізуються.

Результати. Визначено фактори, що впливають на оцінку ризиків, запропоновано використання лінгвістичних змінних для їх опису та використання нечітких змінних для оцінки їх якостей, а також системи якісних оцінок. Обґрунтовано вибір параметрів та розроблено структуру нечіткої продукційної моделі оцінювання ризиків та бази правил нечіткого логічного висновку. Розглянуто використання нечітких моделей для вирішення задач оцінки ризиків інформаційної безпеки, а також концепцію та побудову ERP-систем та проаналізовано проблеми їх безпеки та вразливості.

Висновки. Розроблено нечітку модель оцінки ризиків ERP-системи. Обрано перелік факторів, що впливають на ризик інформаційної безпеки. Запропоновано методи оцінки ризику інформаційних ресурсів та ERP-систем взагалі, оцінки фінансових збитків від реалізації загроз, визначення типу ризику за його оцінкою для формування рекомендацій відносно їх обробки з метою підтримки рівня захищеності ERP-системи. Визначено перелік лінгвістичних змінних моделі. Обрано структуру бази нечітких продукційних правил – MISO-структуру. Побудовано структуру нечіткої моделі. Визначено нечіткі змінні моделі.

КЛЮЧОВІ СЛОВА: інформаційна безпека, нечітка логіка, оцінка ризиків, захищенність, ERB-система.

УДК 004.94

РАЗРАБОТКА НЕЧЕТКОЙ МОДЕЛИ ОЦЕНКИ РИСКОВ ДЛЯ ERP-СИСТЕМЫ

Кожуховский А. Д. – д-р техн. наук, профессор, профессор кафедры информационной та кибернетической безопасности Государственного университета телекоммуникаций, Киев, Украина.

Кожуховская О. А. – д-р техн. наук, доцент кафедры информационной та кибернетической безопасности Государственного университета телекоммуникаций, Киев, Украина.

АННОТАЦИЯ

Актуальность. Поскольку оценка рисков информационной безопасности является сложным и полным процессом неопределенности, а неопределенность является одним из основных факторов, влияющих на эффективность оценки, целесообразно использовать нечеткие методы и модели, которые являются адаптивными к неучтенных данных. Формирование расплывчатых оценок факторов риска субъективно, а оценка рисков зависит от практических результатов, полученных в процессе обработки рисков угроз, которые уже возникли в ходе функционирования организации, и опыта специалистов по информационной безопасности. Поэтому целесообразно использовать модели, которые могут адекватно оценивать нечеткие факторы и иметь возможность корректировать их влияние на оценку рисков. Наибольшими показателями эффективности для решения таких проблем являются нейро-нечеткие модели, сочетающими методы нечеткой логики и искусственные нейронные сети и системы, т.е. «человеко-подобный» стиль соображений нечетких систем с обучением и моделированием психических явлений нейронных сетей. Для построения модели расчета оценки рисков информационной безопасности предлагается использовать нечеткую модель продукта. Нечеткие модели продуктов (нечеткие модели /системы на основе правил) являются обычным типом нечетких моделей, используемых для описания, анализа и моделирования сложных систем и процессов, которые плохо формализованы.

Цель работы – разработка структуры нечеткой модели оценки рисков информационной безопасности и защиты систем ERP с использованием нечетких нейронных моделей.

Метод. Для построения модели расчета оценки рисков информационной безопасности предлагается использовать нечеткую модель продукта. Нечеткие модели продуктов являются обычным видом нечетких моделей, используемых для описания, анализа и моделирования сложных систем и процессов, которые плохо формализованы.

Результаты. Выявленные факторы, влияющие на оценку риска, свидетельствуют об использовании лингвистических переменных для их описания и использования нечетких переменных для оценки их качеств, а также системы качественных оценок. Обоснован выбор параметров и разработана структура нечеткой модели оценки рисков и основы правил нечеткого логического заключения. Рассматривается использование нечетких моделей для решения проблем оценки рисков информационной безопасности, а также концепция и строительство систем ERP и проанализированы проблемы их безопасности и уязвимости.

Выводы. Разработана нечеткая модель оценки рисков системы ERP. Выбран перечень факторов, влияющих на риск информационной безопасности. Предлагаются методы оценки рисков информационных ресурсов и ERP-систем в целом, оценка финансовых потерь от реализации угроз, определение вида риска в соответствии с его оценкой для формирования рекомендаций по их обработке в целях поддержания уровня защиты системы ERP. Определен список лингвистических перемен-

DOI 10. 15588/1607-3274-2022-1-12

ных модели. Выбрана структура базы данных нечетких правил продукта – MISO-структура. Построена структура нечеткой модели. Выявлены нечеткие переменные модели.

КЛЮЧЕВЫЕ СЛОВА: информационная безопасность, нечеткая логика, оценка рисков, защищенность, ERB-система.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Leighton J. Security Controls Evaluation, Testing and Assessment Handbook / J. Leighton. – Syngress, 2016. – 678 p.
2. Rescher N. Many-Valued Logic / N. Rescher. – Mc.Graw-Hill, New York, 1969. DOI:10.2307/2272880
3. Rosser J. B. Many-Valued Logics / J. B. Rosser, A. R. Turquette. – North Holland, Amsterdam, 1952.
4. Common Vulnerability Scoring System version 3.1: Specification Document. CVSS Version 3.1 Release [Електронний ресурс] // Forum of Incident Response and Security Teams. – Режим доступу: https://www.first.org/cvss/ specification-document.
5. Abhishek Kumar Srivastav A Quantitative Measurement Methodology for calculating Risk related to Information Security / Abhishek Kumar Srivastav, Irman Ali, Shani Fatema // IOSR Journal of Computer Engineering (IOSR-JCE). – (Feb. 2014). – Volume 16, Issue 1, Ver. IX. – P. 17–20.
6. Hayashi Y. Fuzzy neural expert system with automated extraction of fuzzy If-Then rules from a trained neural network / Y. Hayashi, A. Imura // Proceedings. First International Symposium on Uncertainty Modeling and Analysis. – 1990. – P. 489–494. DOI.1109/ISUMA.1990.151303
7. Buckleya J. J. Fuzzy neural networks: a survey / J. J. Buckleya, Y. Hayashi // Fuzzy sets and systems. – 1994. – Vol. 66, Issue 1. – P. 1–13. https://doi.org/10.1016/0165-0114(94)90297-6
8. Hendrawirawan D. ERP Security and Segregation of Duties Audit: A Framework for Building an Automated Solution / D. Hendrawirawan, H. Tanriverdi, C. Zetterlund // information systems control journal. – 2007. – Vol. 2. – 4 p. ISACA. All rights reserved. www.isaca.org
9. Nieto-Morote A. A. Fuzzy approach to construction Project risk assessment / A. Nieto-Morote, F. Ruz-Vila // International Journal of Project Management. – 2011. – Vol. 29, Issue 2. – P. 220–231.
10. Kozhukhivskyi A. D. ERP-System Risk Assessment Methods and Models (Tekst) / A. D. Kozhukhivskyi, O. A. Kozhukhivska // Radio Electronics, Computer Science, Control. – 2020. – No. 4(55). – P. 151–162. DOI 10.15588/1607-3274-2020-4-15.
11. Baskerville R. An analysis survey of information system security design methods: Implications for Information Systems Development / R. Baskerville // ACM Computing Survey. – 1993. – P. 375–414.
12. Peltier T. R. Facilitated risk analysis process (FRAP) / T. R. Peltier. – Auerbach Publication, CRC Press LLC, 2000. – 21 p.
13. Alberts C. Managing Information Security Risks: The Octave Approach / C. Alberts, A. Dorofee, Addison-Wesley Professional. – 2002. – 512 p.
14. Stolen K. Model-based risk assessment – the CORAS approach [Elektronnyi resurs] / K Stolen, F. den Braber, T. Dirmitrakos. – 2002. – Rezhim dostupu: http://folk.uio.no/nik/2002/Stolen.pdf
15. Suh B. The IS risk analysis based on business model / B. Suh, I. Han // Information and Management. – 2003. – Vol. 41, No. 2. – P. 149–158.
16. Karabacaka B. ISRAM: Information security risk analysis method / B. Karabacaka, I. Songukpinar. – Computer & Security, March. – 2005. – P. 147–169.
17. Goel S. Information security risk analysis – a matrix-based approach [Elektronnyi resurs] / S. Goel, V. Chen. – University at Albany. – SUNY. – 2005. – Rezhim dostupu: https://www.albany.edu /~goel/publications/goelchen2005.pdf
18. Elky S. An introduction to information system risk management [Elektronnyi resurs] / S. Elky. – SANS Institute InfoSec Reading Room. – 2006. – Rezhim dostupu: https://www.sans.org/reading-room/whitepapers/auditing/introduction-information-system-risk-management-1204.
19. Yazar Z. A. Qualitative risk analysis and management tool – CRAMM [Elektronnyi resurs] / Z. A. Yazar. – SANS Institute InfoSec Reading Room. – 2011. – Rezhim dostupu: https://www.sans.org/reading-room/whitepapers/auditing/qualitative-risk-analysis-management-tool-cramm-83
20. Yurchenko A. G. Building information protection systems on fuzzy sets.Theory and practical solutions / A. G. Yurchenko, A. G. Korchenko. – K. : MK-Press. – 2006. – 320 p.: IL.
21. Security issues in ERP. Security, Audit and Control Features SAP ERP 4th Edition, Audit Program. – Isaca. – 2015. – 574 p.
22. A Complete Guide to the Common Vulnerability Scoring System. Forum of Incident Response and Security Teams (June 2007). Rezhim dostupu: http://www.first.org/cvss/cvss-guide.pdf
23. Polyakov A. ERP Security Deserves Our Attention Now More Than Ever [Elektronnyi resurs] / A. Polyakov // Forbes. – 2017. – Rezhim dostupu: https://www.forbes.com/sites/forbestechcouncil/2017/07/07/erp-security-deserves-our-attention-now-more-than-ever/.
24. NVD Common Vulnerability Scoring System Support v2. National Vulnerability Database. – National Institute of Standards and Technology. Retrieved March 2, 2013.
25. Jang J.-S. R. ANFIS: Adaptive Network – based Fuzzy Inference System / J.-S.R. Jang // IEEE Trans. On Syst. Man and Cybernetics. – 1993. – Vol. 23, No. 3. – P. 665–685.
26. National vulnerability database Release [Elektronnyi resurs] // National Institute of Standards and Technology. – Rezhim dostupu: https://nvd.nist.gov
27. National vulnerability database Release. Vulnerability Metrics [Elektronnyi resurs] // National Institute of Standards and Technology. – Rezhim dostupu: https://nvd.nist.gov/vuln-metrics/cvss.

DOI 10. 15588/1607-3274-2022-1-12