

Харківський національний
університет радіоелектроніки

Kharkiv National
University of Radio Electronics

Державне підприємство
"Південний державний
проектно-конструкторський
та науково-дослідний інститут
авіаційної промисловості"

State Enterprise
"Southern National Design
&
Research Institute
of Aerospace Industries"

**СУЧАСНИЙ СТАН
НАУКОВИХ
ДОСЛІДЖЕНЬ
ТА ТЕХНОЛОГІЙ
В ПРОМИСЛОВОСТІ**

**INNOVATIVE
TECHNOLOGIES
AND
SCIENTIFIC SOLUTIONS
FOR INDUSTRIES**

№ 1 (19), 2022

No. 1 (19), 2022

*Щоквартальний
науковий
журнал*

*Quarterly
scientific
journal*

Харків
2022

Kharkiv
2022

РЕДАКЦІЙНА КОЛЕГІЯ

Головний редактор
Бодянский Євгеній Володимирович,
д-р. техн. наук, професор

Заступник головного редактора
Айзенберг Ігор Наумович,
канд. техн. наук, професор (США);
Шекер Серхат,
д-р. техн. наук, професор (Туреччина)

Члени редколегії:

Артиух Роман Володимирович, канд. техн. наук;
Бабенко Віталіна Олексіївна, д-р. екон. наук, канд. техн. наук, професор;
Безкорований Володимир Валентинович, д-р. техн. наук, професор;
Гасімов Юсіф, д-р. мат. наук, професор (Азербайджан);
Гопсєнко Віктор, д-р. техн. наук, професор (Латвія);
Го Цян, д-р. техн. наук, професор (КНР);
Джавад Хамісабаді, канд. техн. наук, доцент (Іран);
Зайцева Єлена, д-р. техн. наук, професор (Словаччина);
Зачко Олег Богданович, д-р. техн. наук, доцент;
Коваленко Андрій Анатолійович, д-р. техн. наук, професор;
Костін Юрій Дмитрович, д-р. екон. наук, професор;
Левашенко Віталій, д-р. техн. наук, професор (Словаччина);
Лемешко Олександр Віталійович, д-р. техн. наук, професор;
Малєєва Ольга Володимирівна, д-р. техн. наук, професор;
Момот Тетяна Валеріївна, д-р. екон. наук, професор;
Музика Катерина Миколаївна, д-р. техн. наук, професор;
Назарова Галина Валентинівна, д-р. екон. наук, професор;
Невлюдов Ігор Шакирович, д-р. техн. наук, професор;
Опанасюк Анатолій Сергійович, д-р. фіз.-мат. наук, професор;
Павлов Сергій Володимирович, д-р. техн. наук, професор;
Перова Ірина Геннадіївна, д-р. техн. наук, доцент;
Петленков Едуард, канд. техн. наук (Естонія);
Петришин Любомир Богданович, д-р. техн. наук, професор (Польща);
Рубан Ігор Вікторович, д-р. техн. наук, професор;
Семенець Валерій Васильович, д-р. техн. наук, професор;
Сетлак Галина, д-р. техн. наук, професор (Польща);
Терзіян Ваган Якович, д-р. техн. наук, професор (Фінляндія);
Телстов Олександр Сергійович, д-р. екон. наук, професор;
Тимофєєв Володимир Олександрович, д-р. техн. наук, професор;
Філатов Валентин Олександрович, д-р. техн. наук, професор;
Чумаченко Ігор Володимирович, д-р. техн. наук, професор;
Чухрай Наталія Іванівна, д-р. екон. наук, професор;
Юн Джин, канд. фіз.-мат. наук, професор (КНР);
Ястремська Олена Миколаївна, д-р. екон. наук, професор.

ЗАСНОВНИКИ

Харківський національний університет радіоелектроніки,
Державне підприємство "Південний державний
проектно-конструкторський та науково-дослідний
інститут авіаційної промисловості"

АДРЕСА РЕДАКЦІЇ:

Україна, 61166, м. Харків, проспект Науки, 14
Інформаційний сайт: <http://itssi-journal.com>
E-mail редколегії: journal.itssi@gmail.com

EDITORIAL BOARD

Editor in Chief
Bodianskiy Yevgeniy,
Dr. Sc. (Engineering), Professor, Ukraine

Deputy Chief Editor
Igor Aizenberg,
PhD (Computer Science), Professor (United States)
Serhat Seker,
Dr. Sc. (Engineering), Professor (Turkey)

Editorial Board Members:

Artiukh Roman, PhD (Engineering Sciences) (Ukraine);
Babenko Vitalina, Dr. Sc. (Economics); PhD (Engineering Sciences), Professor (Ukraine);
Bezkorovainyi Volodymyr, Dr. Sc. (Engineering), Professor (Ukraine);
Gasimov Yusif, Dr. Sc. (Mathematical), Professor (Azerbaijan);
Gopeyenko Vectors, Dr. Sc. (Engineering), Professor (Latvia);
Guo Qiang, Dr. Sc. (Engineering), Professor (P.R. of China);
Javad Khamisabadi, PhD (Industrial Management), Associate Professor (Iran);
Zaitseva Elena, Dr. Sc. (Engineering), Professor (Slovak Republic);
Zachko Oleh, Dr. Sc. (Engineering), Associate Professor (Ukraine);
Kovalenko Andrey, Dr. Sc. (Engineering), Professor, (Ukraine);
Kostin Yuri, Dr. Sc. (Economics), Professor (Ukraine);
Levashenko Vitaly, Dr. Sc. (Engineering), Professor (Slovakia);
Lemeshko Olexsandr, Dr. Sc. (Engineering), Professor (Ukraine);
Malayeva Olga, Dr. Sc. (Engineering), Professor (Ukraine);
Momot Tetiana, Dr. Sc. (Economics), Professor, (Ukraine);
Muzyka Kateryna, Dr. Sc. (Engineering), Professor (Ukraine);
Nazarova Galina, Dr. Sc. (Economics), Professor (Ukraine);
Nevliudov Igor, Dr. Sc. (Engineering), Professor (Ukraine);
Opanasyuk Anatoliy, Dr. Sc. (Physical and Mathematical), Professor (Ukraine);
Pavlov Sergii, Dr. Sc. (Engineering), Professor (Ukraine);
Perova Iryna, Dr. Sc. (Engineering), Associate Professor (Ukraine);
Petlenkov Eduard, PhD (Engineering Sciences) (Poland);
Petryshyn Lubomyr, Dr. Sc. (Engineering), Professor (Poland);
Ruban Igor, Dr. Sc. (Engineering), Professor, (Ukraine);
Semenets Valery, Dr. Sc. (Engineering), Professor, (Ukraine);
Setlak Galina, Dr. Sc. (Engineering), Professor (Poland);
Terziyan Vagan, Dr. Sc. (Engineering), Professor, (Finland);
Teletov Aleksandr, Dr. Sc. (Economics), Professor (Ukraine);
Timofeyev Volodymyr, Dr. Sc. (Engineering), Professor (Ukraine);
Filatov Valentin, Dr. Sc. (Engineering), Professor (Ukraine);
Chumachenko Igor, Dr. Sc. (Engineering), Professor (Ukraine);
Chukhray Nataliya, Dr. Sc. (Economics), Professor (Ukraine);
Yu Zheng, PhD (Physico-Mathematical Sciences), Professor (P.R. of China);
Iastrenska Olena, Dr. Sc. (Economics), Professor (Ukraine).

ESTABLISHERS

Kharkiv National University of Radio Electronics,
State Enterprise "National Design & Research Institute
of Aerospace Industries"

EDITORIAL OFFICE ADDRESS:

Ukraine, 61166, Kharkiv, Nauka Ave, 14
Information site: <http://itssi-journal.com>
E-mail of the editorial board: journal.itssi@gmail.com

Журнал включено до "Переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата наук" наказом Міністерства освіти і науки України від 16.07.2018 №775 (додаток 7).

Затверджений до друку Науково-технічною Радою Харківського національного університету радіоелектроніки (Протокол № 2 від 30 березня 2022 р.).

Свідоцтво про державну реєстрацію журналу Серія KB № 22696-12596P від 04.05.2017 р.

ЗМІСТ

Інформаційні технології

- 5 **Барковська О. Ю., Хомич В. М., Настенко О. С.**
Дослідження методів обробки та аналізу тексту при організації електронних сховищ інформаційних об'єктів (eng.)
- 13 **Батюк Т. М., Висоцька В. А.**
Розробка інтелектуальної системи підтримки соціалізації користувача за подібністю інтересів
- 27 **Мінухін С. В.**
Дослідження продуктивності моделі DTU для реляційних баз даних на платформі Azure (eng.)
- 40 **Попов А. В., Момот М. О., Єлізєва А. В.**
Вибір системи автоматизації тестування з урахуванням вимог замовника (eng.)
- 47 **Смідович Л. С., Давидовський Ю. К.**
Процеси трансформації інформаційної архітектури оператора зв'язку (eng.)
- 55 **Цао Вейлін, Косенко В. В., Семенов С. Г.**
Дослідження ефективності методу підвищення безпеки програмного забезпечення і обґрунтування практичних рекомендацій з його використання (eng.)

Інженерія та промислові технології

- 65 **Альохіна С. В., Невлюдов І. Ш., Ромашов Ю. В.**
Комп'ютерне моделювання процесів керованості для роботизованих колісних платформ з врахуванням обмежень ривків рухів (eng.)
- 76 **Черняк О. М., Сороколат Н. А., Багасєв І. О., Фатєєва Л. Ю.**
Застосування функціональної залежності для багатокритеріального оцінювання безпеки праці, як об'єкта кваліметрії

Електроніка, телекомунікаційні системи та комп'ютерні мережі

- 85 **Зусєв А. О., Івашко А. В., Лунін Д. О.**
Оцінка програмної складності обчислення коефіцієнтів авторегресії при цифровому спектральному аналізі (eng.)
- 92 **Князєв В. В., Лазуренко Б. О., Серков О. А.**
Методи і засоби оцінки рівня завадостійкості безпроводних каналів зв'язку (eng.)
- 99 **Алфавітний показник**

CONTENTS

Information Technology

- 5 **Barkovska O., Khomych V., Nastenka O.**
Research of the text processing methods in organization of electronic storages of information objects
- 13 **Batiuk T., Vysotska V.**
Intellectual system development for user socialization support by interests similarity
- 27 **Minukhin S.**
Performance study of the DTU model for relational databases on the Azure platform
- 40 **Popov A., Momot M., Yelizieva A.**
Choosing the test automation system according to customer requirements
- 47 **Smidovych L., Davydovskyi Y.**
Processes of the telecom operator's information architecture transformation
- 55 **Cao Weilin, Kosenko V., Semenov S.**
Study of the efficiency of the software security improving method and substantiation of practical recommendations for its use

Engineering & Industrial Technology

- 65 **Alyokhina S., Nevliudov I., Romashov Y.**
Computer simulations of controllability processes for robotic wheeled platforms taking into account restrictions of jerk motions
- 76 **Cherniak O., Sorocolat N., Bahaiev I., Fatieieva L.**
Application of functional dependence for multi-criterial assessment of labor safety as an object of qualimetry

Electronics, Telecommunication Systems & Computer Network

- 85 **Zuev A., Ivashko A., Lunin D.**
Estimation of software complexity of calculation of autoregression coefficients at digital spectral analysis
- 92 **Knyazev V., Lazurenko B., Serkov A.**
Methods and tools for assessing the level of noise immunity of wireless communication channels
- 99 **Alphabetical index**

O. BARKOVSKA, V. KHOMYCH, O. NASTENKO

RESEARCH OF THE TEXT PROCESSING METHODS IN ORGANIZATION OF ELECTRONIC STORAGES OF INFORMATION OBJECTS

The **subject** matter of the article is electronic storage of information objects (IO) ordered by specified rules at the stage of accumulation of qualification thesis and scientific work of the contributors of the offered knowledge exchange system provided to the system in different formats (text, graphic, audio). Classified works of contributors of the system are the ground for organization of thematic rooms for discussion to spread scientific achievements, to adopt new ideas, to exchange knowledge and to look for employers or mentors in different countries. The **goal** of the work is to study the libraries of text processing and analysis to speed-up and increase accuracy of the scanned text documents classification in the process of serialized electronic storage of information objects organization. The following **tasks** are: to study the text processing methods on the basis of the proposed generalized model of the system of classification of scanned documents with the specified location of the block of text processing and analysis; to investigate the statistics of change in the execution time of the developed parallel modification of the methods of the word processing module for the system with shared memory for collections of text documents of different sizes; analyze the results. The **methods** used are the following: parallel digital sorting methods, methods of mathematical statistics, linguistic methods of text analysis. The following **results** were obtained: in the course of the research fulfillment the generalized model of the scanned documents classification system that consist of image processing unit and text processing unit that include unit of the scanned image previous processing; text detection unit; previous text processing; compiling of the frequency dictionary; text proximity detection was offered. **Conclusions:** the proposed parallel modification of the previous text processing unit gives acceleration up to 3,998 times. But, at a very high computational load (collection of 18144 files, about 1100 MB), the resources of an ordinary multiprocessor-based computer with the shared memory obviously is not enough to solve such problems in the mode close to real time.

Keywords: information system; parallelism; word processing; linguistic programming; library; acceleration; method.

Introduction

The paper [1] offers the system of young scientists from different countries knowledge exchange, that provides for communication of the specialists and interested people on the one of available topics. The topic is defined on the basis of qualification thesis and scientific work of the system participants, that prescribes the previous stage of the data collection in database, their analysis to classify and identify topical and scientific and practical directions and authors who are interested in and look into specified directions. Classified works of contributors of the system are the ground for organization of the thematic rooms for discussion to spread scientific achievements, to adopt new ideas, to exchange knowledge and to look for an employers or mentors in different countries.

In the thesis [2] it was offered the organization model of the storage system [3-6] and access to the scientific thesis of the researchers, lecturers and university students within electronic storage IO. Both stages are working with different types of input documents: graphic, text and audio-files (the last format is used only on the data access stage). That is why applicability of the research in the scanned documents classification field for the classification and structured storing in electronic storage of the scientific knowledge exchange system is reasonable.

Analysis of last achievements and publications

Scanned documents processing is a complex process as it represents the combination of the methods of the work with the image and text processing, that was noted in

thesis [7-9]. Further, a lot of works are addressed to the text processing methods [10-14], and realization of methods is performed in numerous linguistic programming libraries – CoreNLP, NLTK, TextBlob, Spacy, Spark NLP etc. The specific feature of a library is a possibility to use different hardware (massive parallel system, general memory systems etc). Processing libraries analysis and text analysis is the goal of the given work. It will provide fast and accurate detected text processing for the further classification.

In the sphere of natural language processing there are many libraries, but the libraries with the limited functionality and operate for the specific cases will not be productive for solving the given task. For instance, NLP gensim library was originally created for thematic modeling and can not be used for complete NLP pipeline construction as a collection of basic and additional functions (fig. 1).

The analysis of the NLP libraries expended features about text processing revealed that the following libraries with the open code meet the necessary requirements: Spark NLP; spaCy; OpenNLP; Stanford CoreNLP., NLTK [15-17]. The comparison of the NLP libraries features is given in table 1.

Although most NLP libraries support learning new models by users it is important that any NLP library was able to provide available pre-trained high-quality models. However, most libraries only support general pre-trained models (POS, NER etc.). Some do not allow the use of their pre-trained models for commercial purposes because of the way they are licensed (table 2).



Fig.1. Definition of basic and additional functions of NLP pipeline

Table 1. The comparison of the NLP libraries features

Name of the text processing method	NLTK	Spark NLP	spaCy	CoreNLP
Sentence detection	+	+	+	+
Tokenization	+	+	+	+
Stemming	+	+	+	+
Lemmatization	+	+	+	+
POS tagging	+	+	+	+
NER	+	+	+	+
Dependency parser	+	+	+	+
Text matcher	–	+	+	+
Date matcher	–	+	–	+
Chunking	+	+	+	+
Spell checking	–	+	–	–
Sentiment analysis	–	+	–	+
Pre-trained models	+	+	+	+
Trained models	+	+	+	+

Table 2. Comparison of NLP libraries with the support of pre-trained models

Name	Pre-trained models (general)	Pre-trained models (subject-oriented)	Free license to use pre-trained models commercially
NLTK	+	–	+
Spark NLP	+	+	+ (partially)
spaCy	+	+	+ (partially)
CoreNLP	+	–	–
OpenNLP	+	–	+

The conducted analysis of the linguistic programming libraries, the functionality of which is focused on natural language processing with ready-made pre-trained models of neural networks, conveyors and

vector representations of words, as well as supporting the learning of their own models, showed that Spark NLP and spaCy are the leaders, therefore, these libraries can be successfully used for research and experimentation in order to reduce the operating time of the text unit in the proposed generalized model of the scanned documents classification system.

Comparing Spark NLP and spaCy, it should be noted that spaCy is the most documented library and possesses an extended training course, as well as the wide industrial application. In the research [18] the above mentioned libraries were analyzed on the accuracy of training, model size, time forecast, F-measures on the same data sets for all libraries in the process of training and testing. The spaCy library provided the most efficient performance and the highest results of the accuracy of training in comparison with the other models. Therefore, the experimental part of this work was performed using the spaCy linguistic library.

Aim of the study is to study the text processing and analysing libraries with the aim to speed up and increase the accuracy of the scanned text documents classifications when organizing an orderly electronic storage of the information objects.

Achieving the goal is possible by solving the following tasks:

- to perform the analysis of the modern linguistic programming libraries for sequential and parallel

implementations of the most common methods of text processing and analysis;

- to offer the generalized model of the scanned documents classification system with the specified place of the block of processing and the analysis of the text;
- to provide a modification of the methods of the pre-processing module for the system with shared memory;
- to obtain statistics of the time spent on the work of the pre-processing module for the collections of the text documents of different sizes;
- to analyze the obtained results.

Materials and methods

The topicality of the research lies in reducing the time of organizing the new information resources entering the storage by increasing the speed of processing methods and analysis of the recognized text in the original image.

Figure 2 shows the place of the methods of text processing and analysis on the generalized model of the scanned documents classification system, which consists of the following units – image and text processing units, which in their turn include the modules of the scanned image pre-processing; text detection; text pre-processing; creating the frequency dictionary; defining the textual proximity. The mentioned units work in the sequence shown in fig. 2, that is the final result (the accuracy and time of classification) depends on all methods used.

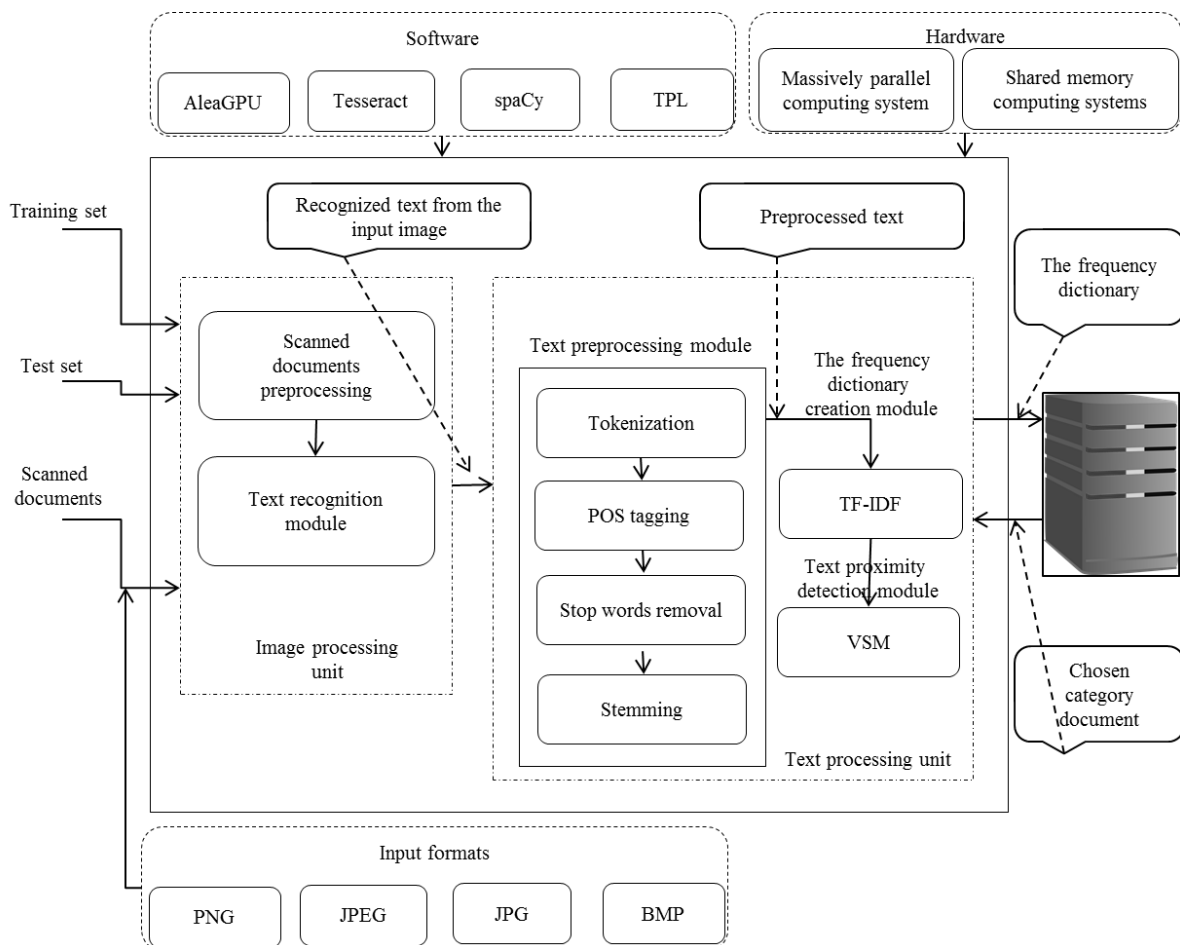


Fig. 2. The scheme of the text processing and analysis methods on the generalized model of the scanned documents classification system

Linguistic pipeline components, represented in the text pre-processing module, are called in order. The tokenizer is launched at the first stage in the NLP-block and as a result leads to breaking a paragraph of the text into smaller parts (tokens), for example, words or sentences.

The original text, on which the components of the conveyor were tested, are given below:

['We will protect our players and I will protect my players always.', 'He played against Everton as a number 6, a number 9 and as a wing-back.', 'Is that fair to him?']

Full breaking of the paragraph into words in the text form is given below:

['We', 'will', 'protect', 'our', 'players', 'and', 'I', 'will', 'protect', 'my', 'players', 'always', '.', 'He', 'played', 'against', 'Everton', 'as', 'a', 'number', '6', ',', 'a', 'number', '9', 'and', 'as', 'a', 'wing-back', '.', 'Is', 'that', 'fair', 'to', 'him', '?']

Removing noise words as such that do not convey any semantic load in the text, for example, "the", "a", "at", "for", "above", "on", "is", "all" is an important step, that influences the accuracy of the original documents classification as well as the speed of the result obtaining by reducing the amount of the text processed. In spaCy in order to remove such words it is necessary to create the noise words list and filter the list of tokens from these words.

The full result of the filtered sentence after removing the noise words from the tokenized sentence is given below in the text form:

['We', 'protect', 'players', 'I', 'protect', 'players', 'always', '.', 'He', 'played', 'Everton', 'number', '6', ',', 'number', '9', 'wing-back', '.', 'Is', 'fair', '?']

Lexical rationing takes into account another noise type in the text. For example, the words "connection", "connected", "connecting" are the derivatives of the general word "connect". Such rationing reduces related word forms to a common root word. The process of shortening a word to the base by discarding auxiliary parts such as an ending or a suffix is stemming. The results of stemming are sometimes very similar to finding the root of a word, but its algorithm is based on other principles. That's why a word after processing by the stemming algorithm can be different from its morphological root.

The result of the filtered sentence stemming (stemmed sentence):

['We', 'protect', 'player', 'I', 'protect', 'player', 'alway', '.', 'He', 'play', 'everton', 'number', '6', ',', 'number', '9', 'wing-back', '.', 'Is', 'fair', '?']

Based on the format of the original data – collection of the text documents – and obtained in the process of their reading string variables with their contents, stream processing of the texts is more effective than processing them one by one in the traditional sequential form.

In this case, one of the documents of the collection is sent for processing by the stream. Testing was done on the system with general memory AMD Ryzen 5 2600 3.4GHz, having 6 cores and a maximum number of streams – 12. SpaCy recommends joblib library for parallel processing of the NLP conveyor units. To

parallelize the workflow, it is necessary to identify a few more auxiliary methods.

Study results and their discussion

In order to analyze the operation of the text processing methods without the use of the parallelization procedure and with the help of it, collections of the documents of different sizes were used (table 3):

- 1 file of 10 thousand words, about 5 MB;
- collection of 25 files of 225 thousand words, about 120 MB;
- collection of 3024 files, about 200 MB;
- collection of 9072 files, about 600 MB;
- collection of 18144 files, about 1100 MB.

The case that was used for testing consists of the files that represent classical works of literature in the plain text format, and in particular:

- A Tale of Two Cities – A Story of the French Revolution;
- Alice's Adventures in Wonderland;
- Fairy Tales by the Brothers Grimm;
- Frankenstein; or, the Modern Prometheus;
- Moby-Dick or, the Whale;
- Pride and Prejudice by Jane Austen;
- The Adventures of Sherlock Holmes;
- The Adventures of Tom Sawyer;
- The Iliad of Homer;
- The Romance of Lust by Unknown writer;
- War and Peace by Leo Tolstoy etc.

For each of the methods listed in the word processing module testing was conducted on each of the five collections. The testing included 10 application launches to determine the average execution time of the methods.

Fig. 3 represents the block diagram with parallel modification of the methods of the pre-processing module for the system with shared memory.

For the systems with shares memory the programme model for the decomposition of the traditional sequential approach is the fork-join model, which involves generation of parallel flows, their independent performance, combination of the results in the main stream. Thus, the operating time of the module is significantly reduced and instead of the sum of the processing times of all documents the area of preparation of the original text for the creation of the frequency dictionary is equal to the time of the longest processing, which corresponds to the most voluminous original text.

The time, given in the table in the line without acceleration, is the sum of the time of tokenization methods, removal of stop words, POS tagging and steaming. Here is an example for a file of 10 thousand words: the time of tokenization is 0,054 seconds, the time of the stop words removal is 0,012 seconds, the time for POS tagging is 0,583 seconds and the time for stemming is 0,1524 seconds. That is, the total operating time of the pre-processing module without acceleration and use of nlp.pipe is 1,3964 seconds for the collection including only one file with the volume of 10 thousand words.

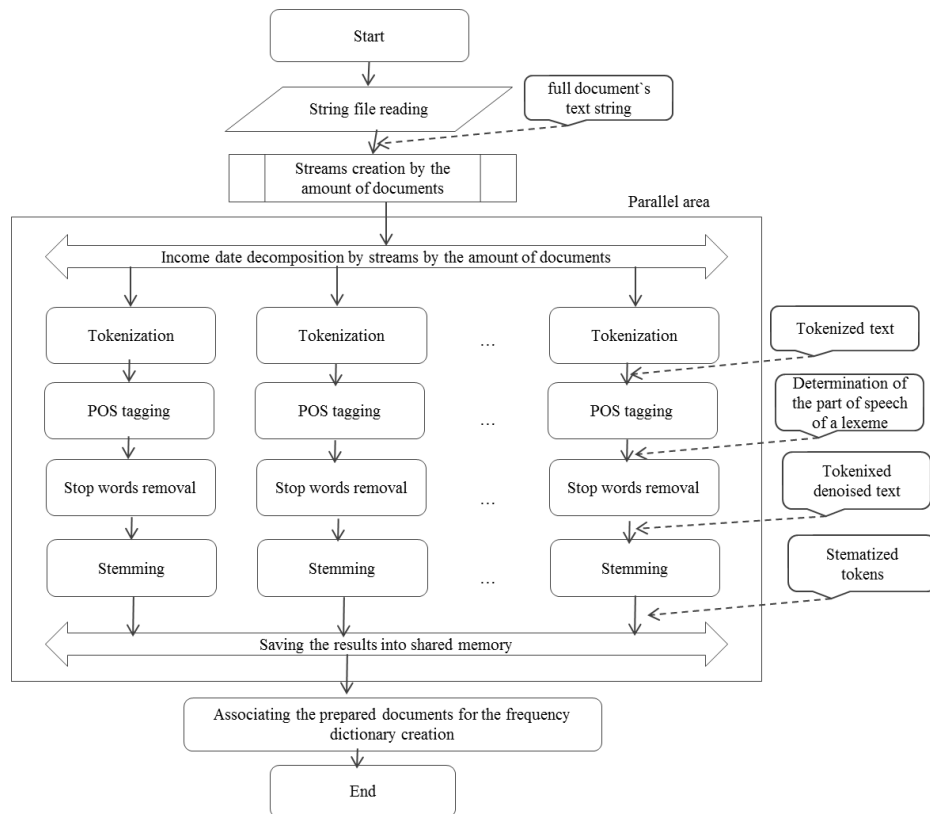


Fig.3. Block diagram with parallel modification of the preprocessing methods for the shared memory systems

Table 3. Statistics of the time spent on the operation of the text pre-processing module

Method		Number of files				
		1	25	3024	9072	18144
Without acceleration		1,3964 sec	8,42 sec	758,7 sec	1956,78 sec	5894,6 sec
nlp.pipe		1,781 c	6,354c	620,35c	1604,9c	5691,4c
Paralleling	2 streams	1,413 c	3,247c	350,4c	821,3c	3854,8M
	4 streams	1,348 c	2,87c	192,7c	443,46c	2310,7c

According to the results depicted on the graph (fig. 4), we can observe the tendency of acceleration but the resources of an ordinary multiprocessor-based computer with the shared memory obviously is not enough to solve such problems in the mode close to real time. Parallel

algorithm for several streams has an advantage only in case of presence of a large number of documents, and the higher the number of streams the greater the acceleration. If it is necessary to analyze a small amount of text it is better to divide the data into parts and transfer to nlp.pipe.

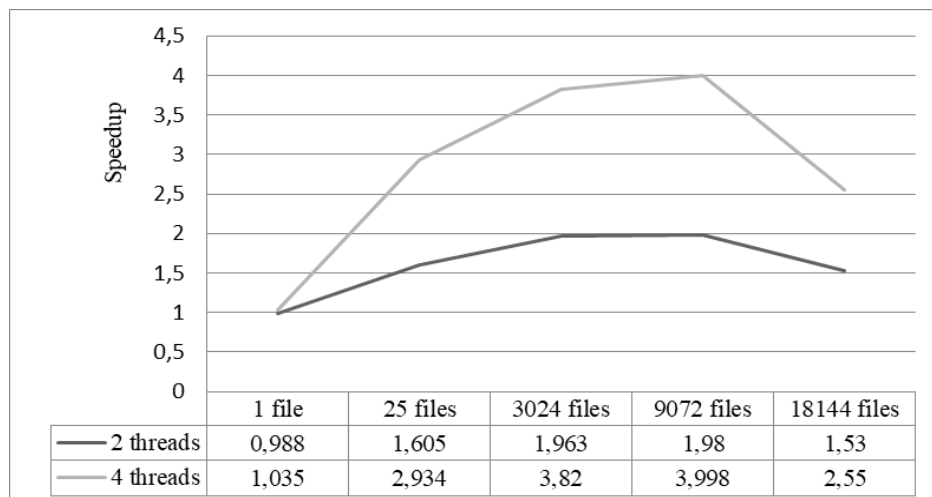


Fig.4. Graph of the text preliminary processing methods acceleration

The next step of the functioning of the text processing unit is the work on the creation of the frequency dictionary module and the module of determining the textual proximity. As the compilation and organization of the frequency dictionary is given in the article of the authors of this survey [2], further research will be related exactly to the definition of the textual proximity which is an important step for the efficient storage of documents and its organization. Determining the percentage of similarity between the texts in the given generalized classification system of the scanned documents is proposed to avoid duplication of documents, that is resaving of an existing document with a different name.

Conclusion

In the course of the research it was proposed a generalized classification system of the scanned documents, which consists of the image and text processing units, which in their turn include the modules of the scanned image pre-processing; text detection; text pre-processing; creating the frequency dictionary; defining the textual proximity. As the given units operate in the sequence, that is the final result (the accuracy and time of classification) depends on all methods used, the reduction of the time for classification and arrangement of the new information resources entering the system is possible by increasing the speed of processing methods and analysis of the recognized text in the original image.

Therefore, in the research it was proposed the parallel modification of the methods of the pre-processing module for the system with the shared memory and conducted its software testing with the definition of the obtained acceleration depending on the different number of the computational streams used.

Analysis of the libraries of linguistic programming, the functionality of which is focused on the processing of natural language with ready-made pre-trained models of neural networks, conveyors and vector representations of the words as well as the support of learning their own models, showed that Spark NLP and spaCy are the leaders, although, such advantages as the best documentation, industrial application, high productivity and accuracy of learning compared to other libraries determined the experimental research using the spaCy linguistic programming library.

The results showed a tendency to accelerate up to 3,998 times but, at a very high computational load (collection of 18144 files, about 1100 MB), the resources of an ordinary multiprocessor-based computer with the shared memory obviously is not enough to solve such problems in the mode close to real time. If it is necessary to analyze a small amount of text it is recommended to divide the data into parts and transfer to nlp.pipe.

Perspectives of further development

Future research will be further concentrated on reducing the time required to prepare text documents for classification, focusing the attention on large collections (more than 200 thousand words). Besides, significant is the time of the method of determining the textual proximity, which is an important step for efficient storage and organization of documents.

To achieve greater acceleration, it is planned to conduct the experiments with the application of the mass parallelism computer systems and various approaches to the original data decomposing – ordinal distribution, sentence distribution, adaptive distribution.

References

1. Barkovska, O., Kholiev, V., Pyvovarova, D., Ivaschenko, G., Rosinskiy, D. (2021), "International system of knowledge exchange for young scientists", *Advanced Information Systems*, No. 5 (1), P. 69 – 74. DOI: <https://doi.org/10.20998/2522-9052.2021.1.09>
2. Barkovska, O., Pyvovarova, D., Kholiev, V., Ivashchenko, H., Rosinskiy, D. (2021), "Information Object Storage Model with Accelerated Text Processing Methods", *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021)*, No. 2870, P. 286 – 299.
3. Koroteev, M. (2020), "On the Usage of Semantic Text-Similarity Metrics for Natural Language Processing in Russian", *13th International Conference "Management of large-scale system development" (MLSD)*, P. 1 – 4. DOI: <https://doi.org/10.1109/MLSD49919.2020.9247691>
4. Liu, Y. Sheng, Wei, Z., Yang, Y. (2018), "Research of Text Classification Based on Improved TF-IDF Algorithm", *IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE)*, P. 218 – 222. DOI: <https://doi.org/10.1109/IRCE.2018.8492945>
5. Zhang, Y. (2021), "Research on Text Classification Method Based on LSTM Neural Network Model", *IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, P. 1019 – 1022. DOI: <https://doi.org/10.1109/IPEC51340.2021.9421225>
6. Jindal, R., Shweta, (2018), "A Novel Method for Efficient Multi-Label Text Categorization of research articles", *International Conference on Computing, Power and Communication Technologies (GUCON)*, P. 333 – 336. DOI: <https://doi.org/10.1109/GUCON.2018.8674985>
7. Martínek, J., Lenc, L., Král, P. (2020), "Building an efficient OCR system for historical documents with little training data", *Neural Computing and Applications*, No. 32, P. 17209 – 17227. DOI: <https://doi.org/10.1007/s00521-020-04910-x>
8. Pawar, N., Shaikh, Z., Shinde, P., Warke Y. (2019), "Image to Text Conversion Using Tesseract", *International Research Journal of Engineering and Technology (IRJET)*, No. 6 (2), P. 516– 519.
9. Revathi, A., Modi, N. A. (2021), "Comparative Analysis of Text Extraction from Color Images using Tesseract and OpenCV", *8th International Conference on Computing for Sustainable Global Development (INDIACom)*, P. 931 – 936. DOI: <https://doi.org/10.1109/INDIACom51348.2021.00167>

10. Burns, S. (2019), *Natural Language Processing: A Quick Introduction to NLP with Python and NLTK (Step-by-Step Tutorial for Beginners)*, Amazon KDP Printing and Publishing C, 123 p.
11. Lane, H., Hapke, H., Howard, C. (2019), *Natural Language Processing in Action: Understanding, analyzing, and generating text with Python*, Manning; 1st edition, 544 p.
12. Jurafsky, D., Martin, J.H., "Speech and Language Processing", available at: <https://web.stanford.edu/~jurafsky/slp3/> (last accessed: 16.02.2022)
13. Kim, J., Hur, S., Lee, E., Lee, S. (2021), "NLP-Fast: A Fast, Scalable, and Flexible System to Accelerate Large-Scale Heterogeneous NLP Models," *30th International Conference on Parallel Architectures and Compilation Techniques (PACT)*, P. 75– 89. DOI: <https://doi.org/10.1109/PACT52795.2021.00013>
14. Berko, A., Matseliukh, Y., Ivaniv, Y., Chyrun, L., Schuchmann, V. (2021), "The Text Classification Based on Big Data Analysis for Keyword Definition Using Stemming," *IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)*, P. 184– 188. DOI: <https://doi.org/10.1109/CSIT52700.2021.9648764>
15. Sakthi vel, S. (2021), "Pre-Processing techniques of Text Mining using Computational Linguistics and Python Libraries," *International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, P. 879–884. DOI: <https://doi.org/10.1109/ICAIS50930.2021.9395924>
16. Al Omran, F. N. A., Treude, C. (2017), "Choosing an NLP Library for Analyzing Software Documentation: A Systematic Literature Review and a Series of Experiments," *IEEE/ACM 14th International Conference on Mining Software Repositories (MSR)*, P. 187– 197. DOI: <https://doi.org/10.1109/MSR.2017.42>.
17. Vasiliev, Y. (2020), *Natural Language Processing with Python and SpaCy: A Practical Introduction*, No Starch Press, 217 p
18. Naseer, S., Mudasar Ghafoor, M., Alvi, S. bin K., Kiran, A., Shafique Ur Rahmand, Ghulam Murtazae, & Murtaza, G. (2022), "Named Entity Recognition (NER) in NLP Techniques, Tools Accuracy and Performance", *Pakistan Journal of Multidisciplinary Research*, No. 2 (2), P. 293– 308.

Received 25.02.2022

Відомості про авторів / Сведения об авторах / About the Authors

Барковська Олеся Юрївна – кандидат технічних наук, доцент, доцент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна; e-mail: olesia.barkovska@nure.ua; ORCID ID: <https://orcid.org/0000-0001-7496-4353>.

Барковская Олеся Юрьевна – кандидат технических наук, доцент, доцент кафедры Электронных вычислительных машин, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

Barkovska Olesia – Ph.D (Engineering Sciences), Docent, Associate Professor Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Хомич Віктор Михайлович – студент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна; e-mail: viktor.khomych@nure.ua.

Хомич Виктор Михайлович – студент кафедры Электронных вычислительных машин, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

Khomych Viktor – students of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Настенко Олександр Сергійович – студент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна; e-mail: oleksandr.nastenko2@nure.ua.

Настенко Александр Сергеевич – студент кафедры Электронных вычислительных машин, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

Nastenko Oleksandr – students of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

ДОСЛІДЖЕННЯ МЕТОДІВ ОБРОБКИ ТА АНАЛІЗУ ТЕКСТУ ПРИ ОРГАНІЗАЦІЇ ЕЛЕКТРОННИХ СХОВИЩ ІНФОРМАЦІЙНИХ ОБ'ЄКТІВ

Предметом дослідження в статті є електронне сховище інформаційних об'єктів, впорядковане за визначеними правилами на етапі накопичення кваліфікаційних та наукових робіт учасників запропонованої системи обміну знаннями, наданими до системи у різних форматах (текстові, графічні, аудіо). Класифіковані роботи учасників системи є підставою для організації тематичних кімнат для обговорення із метою розповсюдження наукових досягнень, запозичення нових ідей, обміну знаннями та пошуку роботодавців або менторів у різних країнах світу. **Мета** роботи – дослідження бібліотек обробки та аналізу тексту із метою прискорення та збільшення точності класифікації сканованих текстових документів при організації впорядкованого електронного сховища інформаційних об'єктів. В статті вирішуються наступні **завдання**: дослідити методи обробки та аналізу тексту на базі запропонованої узагальненої моделі системи класифікації сканованих документів із зазначеним місцем блоку обробки та аналізу тексту; дослідити статистику зміну часу виконання розробленої паралельної модифікації методів модулю попередньої обробки тексту для системи із загальною пам'яттю для колекцій текстових документів різного розміру; проаналізувати отримані результати. Використовуються такі **методи**: паралельні чисельні методи сортування, методи

математичної статистики, лінгвістичні методи аналізу тексту. Отримано наступні **результати**: в ході виконання досліджень, було запропоновано узагальнену модель системи класифікації сканованих документів, яка складається з блоку роботи із зображенням та блоку роботи із текстом, які, в свою чергу, включають модулі попередньої обробки сканованого зображення; модуль розпізнавання тексту; попередньої обробки тексту; побудови частотного словнику; визначення текстової близькості. **Висновки**: запропонована паралельна модифікація модулю попередньої обробки тексту дає прискорення до 3,998 разів. Але, при дуже високому обчислювальному навантаженні (колекція з 18144 файлів, близько 1100Мб), ресурсів обчислювача на базі багатопроцесорного ЦПУ із загальною пам'яттю не достатньо для вирішення подібних задач у режимі, наближеному до реального часу.

Ключові слова: інформаційна система; паралелізм; обробка тексту; лінгвістичне програмування; бібліотека; прискорення; метод.

ИССЛЕДОВАНИЕ МЕТОДОВ ОБРАБОТКИ И АНАЛИЗА ТЕКСТА ПРИ ОРГАНИЗАЦИИ ЭЛЕКТРОННЫХ ХРАНИЛИЩ ИНФОРМАЦИОННЫХ ОБЪЕКТОВ

Предметом исследования в статье является электронное хранилище информационных объектов, упорядоченное по определенным правилам на этапе накопления квалификационных и научных работ участников предлагаемой системы обмена знаниями, поступающими в систему в различных форматах (текстовые, графические, аудио). Классифицированные работы участников системы являются основанием для организации тематических комнат для обсуждения с целью распространения научных достижений, заимствования новых идей, обмена знаниями и поиска работодателей или менторов в разных странах мира. **Цель** работы – исследование библиотек обработки и анализа текста с целью ускорения и увеличения точности классификации сканированных текстовых документов при организации упорядоченного электронного хранилища информационных объектов. В статье решаются следующие **задачи**: исследовать методы обработки и анализа текста на основе предложенной обобщенной модели системы классификации сканированных документов с указанным местом блока обработки и анализа текста; исследовать статистику изменения времени выполнения разработанной параллельной модификации методов модуля предварительной обработки текста для системы с общей памятью для коллекций текстовых документов разного размера; проанализировать полученные результаты. Используются следующие **методы**: параллельные численные методы сортировки, методы математической статистики, лингвистические методы анализа текста. Получены следующие **результаты**: в ходе выполнения исследований была предложена обобщенная модель системы классификации сканированных документов, состоящая из блока работы с изображением и блока работы с текстом, которые, в свою очередь, включают модули предварительной обработки сканируемого изображения; модуль распознавания текста; предварительной обработки текста; построения частотного словаря; определение текстовой близости. **Выводы**: предложенная параллельная модификация модуля предварительной обработки текста дает ускорение в 3,998 раза. Но, при очень высокой вычислительной нагрузке (коллекция из 18144 файлов, около 1100Мб), ресурсов вычислителя на базе многопроцессорного ЦПУ с общей памятью недостаточно для решения подобных задач в режиме, приближенном к реальному времени.

Ключевые слова: информационная система; параллелизм; обработка текста; лингвистическое программирование; библиотека; ускорение; метод.

Бібліографічні описи / Bibliographic descriptions

Барковська О. Ю., Хомич В. М., Настенко О. С. Дослідження методів обробки та аналізу тексту при організації електронних сховищ інформаційних об'єктів. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 5–12. DOI: <https://doi.org/10.30837/ITSSI.2022.19.005>

Barkovska, O., Khomych, V., Nastenko, O. (2022), "Research of the text processing methods in organization of electronic storages of information objects", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 5–12. DOI: <https://doi.org/10.30837/ITSSI.2022.19.005>

Т. М. БАТЮК, В. А. ВИСОЦЬКА

РОЗРОБКА ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ПІДТРИМКИ СОЦІАЛІЗАЦІЇ КОРИСТУВАЧА ЗА ПОДІБНІСТЮ ІНТЕРЕСІВ

Об'єктом дослідження є процес соціалізації особистостей, оскільки в наш час завдання соціалізації є дуже важливим і всі сучасні соціальні мережі намагаються максимально оптимізувати та автоматизувати соціалізацію різноманітних користувачів з використанням популярних сучасних ІТ, таких як нейронні мережі та алгоритми аналізу користувацького контенту. **Предметом** дослідження є методи та технологічні засоби пошуку та формування списку релевантних користувачів за подібністю інтересів для здійснення соціалізації. Відповідно здійснюється аналіз профіля користувача системи, а саме визначення справжності користувача на основі пошуку людського обличчя на користувацьких фотографіях з використанням нейронних мереж і аналізу користувацького контенту через алгоритми нечіткого пошуку та моделі Noisy Channel. **Мета** роботи – створення інтелектуальної системи для соціалізації особистостей на основі нечіткого пошуку за ключовими словами, моделі Noisy Channel, алгоритму ефективного розподілу текстового контенту та згорткової нейронної мережі. В статті вирішуються наступні **завдання**: проаналізувати сучасні та найвідоміші підходи, методи, засоби та алгоритми розв'язування задач соціалізації особистостей за подібністю інтересів; розробка загальної структури типової інтелектуальної системи соціалізації особистостей за спільними інтересами; сформулювати функціональні вимоги до основних модулів структури типової інтелектуальної системи соціалізації особистостей за спільними інтересами; розробити інтелектуальну систему підтримки соціалізації користувача за подібністю інтересів на основі нейронних мереж, нечіткого пошуку та моделі Noisy Channel та провести експериментальну апробацію. Використовуються такі **методи**: метод Левенштейна; алгоритм N-грам; модель Noisy Channel; нечіткий пошук. Отримано наступні **результати**: побудовано та описано загальну структуру типової інтелектуальної системи соціалізації особистостей за спільними інтересами. До основного призначення створення системи в першу чергу варто віднести створення нового алгоритму аналізу користувацького контенту та пошуку найбільш релевантних за інтересами користувачів на основі таких алгоритмів як N-грам, розширення вибірки, Левенштейна, та моделі Noisy Channel. Також зазнав подальшого розвитку шаблон асинхронного створення програмного продукту, що дозволяє створити практично повністю динамічну систему. Варто удосконалити згорткову нейронну мережу, що дозволить ефективно та динамічно здійснювати пошук людських обличчя на фотографії, та перевіряти наявність вже існуючих людей в базі даних системи. **Висновки**: виявлено, що реалізований алгоритм здійснює вибірку приблизно в 10 разів швидше, ніж звичайний алгоритм Левенштейна. Також реалізований в системі алгоритм для формування вибірки користувачів є ефективнішим та точнішим приблизно на 25-30% в порівнянні зі звичайним алгоритмом Левенштейна.

Ключові слова: відстань Левенштейна; алгоритм N-грам; модель Noisy Channel; нечіткий пошук.

Вступ

Розроблення та впровадження інтелектуальної системи (ІС) соціалізації особистостей є актуальною та перспективною задачею, так як сьогодні користувачі соціальних мереж намагаються оптимізувати життєві процеси для економії часу і відповідно, для більш доцільної/ефективної його реалізації, зокрема, для автоматизації пошуку інших подібних до власних інтересів користувачів соціальних мереж. Така ІС вирішує дві важливі проблеми теперішнього часу інформатизації та глобалізації: автоматизація процесу соціалізації користувачів та максимально оптимізує економію часу для організації буденних потреб у спілкуванні та корисному проведенні часу з однодумцями. Створення ІС ефективного аналізу та підбору користувачів є важливим завданням, так як на даний момент не існує подібних систем. Система соціалізації дозволить оптимізувати процес пошуку та знайомства людей в інтернеті. Для реалізації даної ІС потрібно використати згорткову нейронну мережу для пошуку обличчя на фотографії. Також необхідно застосувати алгоритми нечіткого пошуку та модель Noisy Channel для аналізу інформації про користувача і формування списку релевантних користувачів. Для успішного створення ІС соціалізації особистостей за спільними інтересами найважливішим завданням є правильно зрозуміти та опрацювати процес соціалізації

користувачів.

Аналіз останніх досліджень і публікацій

В напрямку дослідження та реалізації процесів соціалізації сьогодні існує досить велика кількість публікацій, зокрема, в [1-2] автори пропонують нову систему рангів серед користувачів соціальної мережі та адаптивну мережеву модель для зручної взаємодії між користувачами. Також автори пропонують удосконалити систему агрегації та сортування інформації в соціальних мережах відповідно до заданих параметрів. В [3-4] автори пропонують удосконалити можливості соціальних мереж в якості аналізу даних користувачів, формування характеристик роботи користувача та удосконалити можливості соціальних мереж через пошук релевантної подібності між користувачами та, в подальшому, визначення основних інтересів користувача і видача йому найбільш релевантних співпадінь при кожному наступному пошуку контенту в соціальній мережі. В [5] автори пропонують новий підхід для сортування/зберігання даних в хмарі та їх видачу при пошуку з використанням нової гнучкої моделі вишукування даних в соціальних мережах. Основною перевагою подібної ІС є те, що опису практично аналогів або немає (Tinder та Badoo) [6], або це є закритими комерційними проектами, тому категорично стверджувати, що аналоги існують є не припустимим. Подібність ІС з Tinder та Badoo є в

застосуванні згорткової нейронної мережі, яка перевіряє наявність аналогічних користувачів та використовується для ідентифікації користувачів системи [7]. Основним недоліком Tinder та Badoo є те, що вони надають максимально обмежений механізм соціалізації, дозволяючи відфільтрувати користувачів по статі, віку та місцезнаходженню, що є неоптимізованим процесом і несуттєво економить витрачений час на соціалізацію [8].

Виділення не вирішених раніше частин загальної проблеми. Мета роботи

Метою даної статті є створення ІС соціалізації особистостей на основі нечіткого пошуку за ключовими словами на базі моделі Noisy Channel з алгоритмами ефективного розподілу контенту [9] та згорткової нейронної мережі для ідентифікації користувачів. На сьогодні не існує подібної ІС аналізу користувацького контенту та формування списку найбільш подібних за інтересами користувачів. Об'єктом аналізу даних для соціалізації є персональні профілі потенційних релевантних користувачів [10] соціальних мереж. ІС має аналізувати інформацію про користувачів і на основі цього формувати посортований список користувачів за вирахуванням відсотком подібності [11] між користувачами ІС, що і є основною перевагою системи в порівнянні з її основними аналогами [12]. ІС повинна складатися з таких етапів: реєстрація/авторизація користувача, ознайомлення користувача з правилами користування системою [13], завантаження фотографії (на основі згорткової нейронної мережі перевірено на наявність людського обличчя на фото та у вже існуючій базі даних ІС) [14], додавання аватару користувача та заповнення альбому фотографіями. Наступним етапом є введення користувачем контенту про себе, аналіз вказаної інформації та її порівняння з контентом про вже існуючих в системі користувачів з використанням алгоритму Левенштейна, алгоритмів нечіткого пошуку та моделі Noisy Channel, після чого буде сформований посортований список користувачів за вирахуванням відсотком подібності між користувачами. Є можливість додатково відфільтрувати сформований список, змінити особисту інформацію в результаті чого буде наново сформований список, або вийти в головне меню системи. Також всі зміни, внесені користувачем, та переписки зберігатимуться в БД, а всі фотографії в хмарному сервісі збереження медіа-даних. Останнім кроком є дослідження профілю та інтересів [15] обраного з сформованої вибірки користувача та початок спілкування з обраним користувачем, вся переписка у вигляді окремого потоку повідомлень буде зберігатися в БД.

Якщо здійснити постановку та обґрунтування проблеми, то її можна поділити на дві основні частини, а саме реалізація роботи двох нейронних мереж: згорткової та сіамської, що дозволить здійснити пошук людського обличчя та відтворення алгоритмів SEO-технологій, а саме алгоритму

Левенштейна та моделі Noisy Channel, які необхідні для здійснення соціалізації користувачів за спільними інтересами всередині ІС. В наш час важливість використання машинного навчання в сучасних ІС соціалізації особистостей є безперечно важливою, так як всі соціальні мережі оперують великою кількістю даних [16], часто необхідно здійснити розподіл даних, що дозволить оптимізувати та пришвидшити роботу системних процесів, цього можна досягти здійснивши обробку частини даних з використанням машинного навчання. Важливо також не забути про особливості створення серверних програм, а саме оптимізацію алгоритмів взаємодії серверних програм, реалізувати алгоритм оптимізованої пакетної передачі даних різних розмірів. Необхідно реалізувати алгоритми аналізу текстових даних в соціальних мережах [17], створити нейронну мережу, яка порівнює фотографії користувача та текстову інформацію залишену користувачем і на основі проаналізованих даних формує вибірку. Важливо реалізувати оптимальне збереження даних на мобільних пристроях, сформувати основні шаблони збереження обробки та формування вибірки даних з використанням поточної вибірки. Потрібно оптимізувати алгоритми пошуку інформації та здійснити реалізацію механізмів видобування даних, створити нейронну мережу, яка аналізує на вході всю подану соціальну інформацію та з використанням вхідних параметрів відкидає нерелевантну інформацію, зберігаючи лише необхідну в певний момент вибірку даних. Також потрібно не забувати здійснювати оптимізацію процесів всередині системи, а саме зберігання даних на серверах системи та в локальній БД пристрою, та оптимізацію алгоритмів передачі та зберігання даних.

Функціонал системи

Одним з основних завдань є реалізація алгоритму Левенштейна, оскільки він виконує основну роботу ІС, зокрема, пошук відстаней між словами, тобто аналізує подібність слів в контенті. Основна суть алгоритму Левенштейна полягає в тому, що можна побудувати кінцевий автомат, який точно розпізнає набір рядків на заданій відстані Левенштейна до цільового слова. Потім можливо подати будь-яке слово, і автомат прийме або відхилить його залежно від того, чи є відстань Левенштейна до цільового слова щонайбільше тією відстанню, яка наперед визначена, що дозволяє використовувати цей алгоритм для великих обсягів текстового контенту. Крім того, оскільки створюється кінцевий автомат, він здійснює всю роботу за $O(n)$ часу з довжиною рядка, що тестується. Можна зробити висновок, що даний алгоритм ідеально підходить для роботи з середніми об'ємами тексту, оскільки виконується лінійно і швидко здійснює обробку підготовленої текстового контенту.

Ініціалізується матриця, яка вимірює (m, n) - комірки відстані Левенштейна між m -символьним префіксом одного з n -префіксом іншого слова. Матрицю заповнюють від верхнього лівого до

нижнього правого кута. Кожен стрибок по горизонталі або вертикалі відповідає вставці/видаленню відповідно. Вартість зазвичай встановлюється на 1 для кожної з операцій. Діагональний стрибок може коштувати 1-2 символи, коли і в рядку і в стовпці символи не збігаються, або 0, якщо збігаються. Кожна

клітинка завжди мінімізує вартість локально. Таким чином число в нижньому правому куті є отримана відстань Левенштейна між обома словами. В табл. 1 подано приклад аналізу двох слів та знаходження відстаней між ними, яка в результаті дорівнює значенню 4.

Таблиця 1. Приклад роботи алгоритму Левенштейна

words	w1	m	e	i	l	e	n	s	t	e	i	n
w2	0	1	2	3	4	5	6	7	8	9	10	11
l	1	1	2	3	3	4	5	6	7	8	9	10
e	2	2	1	2	3	3	4	5	6	7	8	9
v	3	3	2	2	3	4	4	5	6	7	8	9
e	4	4	3	3	3	3	4	5	6	6	7	8
n	5	5	4	4	4	4	3	4	5	6	7	7
s	6	6	5	5	5	5	4	3	4	5	6	7
h	7	7	6	6	6	6	5	4	4	5	6	7
t	8	8	7	7	7	7	6	5	4	5	6	7
e	9	9	8	8	8	7	7	6	5	4	5	6
i	10	10	9	8	9	8	8	7	6	5	4	5
n	11	11	10	9	9	9	8	8	7	6	5	4

Модель Noisy Channel є алгоритмом для визначення правильного слова в ситуаціях, коли воно незрозуміле після виконання алгоритму Левенштейна. Шум у цьому контексті – це все, що приховує сигнали та дані. Модель зашумленого каналу отримала таку назву, тому що вихідний сигнал – призначене слово – приховане під час передачі, коли порушення або помилки створюють шум у каналі. У цьому випадку шум полягає в орфографічній помилці. Якщо слово не знайдено в словнику перевірки орфографії, воно ідентифікується як орфографічна помилка, і слова-кандидати пропонуються на основі їхньої ймовірності бути призначеним словом, як правило, в результаті того, наскільки ці слова близькі до орфографічної помилки. Як правило, найімовірніші кандидати включатимуть одну зміну, і є чотири різні типи помилок в співставленні літер: видалення, вставка, заміна та скасування. Найпростішим прикладом однієї зміни може бути: відсутні літери, наприклад, leter замість letter, написано зайві літери, наприклад, misstake, замість mistake, літери переставлені, наприклад, recieved замість received, замінені літери, наприклад, fimate замість finite.

Алгоритм N-грам широко використовується, є досить простий та високоефективний. Ідея методу полягає в тому, що якщо два слова A та B збігаються з огляду на можливі помилки та неточності, вони, ймовірно, матимуть спільні n-грами, де n – довжина підрядка. Під час індексації слово чітко ділиться на N-грами, а потім це слово ввійде до словника для кожної N-грами. Під час пошуку, поточний запит також ділиться на N-грами, а для кожної з них виконується послідовний пошук списку слів, що містять такий підрядок. Найбільш часто використовуваним на практиці є 3-грами – підрядки з довжиною в 3 символи. Вибір більшого значення n

призводить до обмеження мінімальної довжини слова, при якій помилки вже можуть бути виявлені.

ІС соціалізації особистостей за спільними інтересами на основі SEO-технологій та методів машинного навчання має одну зовнішню сутність, а саме: *Користувач* і основні внутрішні сутності, до яких належать: *Браузер*, *Контролер системи*, *База даних*, *Сесія користувача*, *Сервер* та *Сервіс системи*. Під час функціонування ІС соціалізації особистостей здійснюється постійна взаємодія між зовнішніми і внутрішніми сутностями:

1. При ініціалізації роботи ІС зовнішня сутність *Користувач* входить в систему за допомогою внутрішньої сутності – *Браузер*.

2. За допомогою сутності *Браузер* та внутрішньої сутності *Контролер системи* здійснюється авторизація або реєстрація користувача в ІС.

3. Після успішної реєстрації/авторизації сутність *Контролер системи* надсилає сесійний токен в сутність *Браузер* та зберігає поточний токен всередині браузера на час існування сесії.

4. Сутність *Користувач* здійснює завантаження власної фотографії для підтвердження справжності за допомогою сутності *Браузер*, який додає фотографію за допомогою сутності *Контролер системи*.

5. Сутність *Контролер системи* на основі нейронних мереж перевіряє наявність обличчя на фотографії сутності *Користувач* та перевіряє наявність вже аналогічних фотографій в БД ІС.

6. Сутність *Контролер системи* на основі сутності *Сервер* зберігає всі отримані дані про користувача та надсилає сформовані дані в іншу внутрішню сутність *База даних*.

7. Сутність *База даних* зберігає всю інформацію на допомогу внутрішньої сутності *Сервіс системи*.

8. *Сервіс системи* опрацьовує інформацію та створює список унікальних користувачів системи.

9. *Сервіс системи* на основі алгоритмів нечіткого пошуку здійснює аналіз даних користувача, який в даний момент працює всередині системи та інших зареєстрованих користувачів.

10. На основі результаті роботи алгоритмів нечіткого пошуку сутностей *Сервіс системи* та *Сесія користувача* створює список користувачів, сформований по спаданню відсоткової подібності між ними відповідно до даних поточного користувача.

11. Сутність *Сервіс системи* здійснює кінцеву перевірку отриманих даних на наявність пошкоджених та недостовірних даних та зберігає отриману інформацію за допомогою сутності *База даних*.

12. За допомогою користувацького запиту сформовані дані з сутності *База даних* надсилаються сутності *Користувач*.

13. *Користувач* здійснює перегляд наданої інформації та вибирає профіль користувача, який йому найбільше сподобався з використанням сутності *Сервіс системи*.

14. Сутність *Користувач* здійснює дію проставлення вподобання іншому користувачу, для того щоб інший користувач системи міг побачити, що його обрав поточний користувач.

15. Сутності *Сервіс системи* та *Контролер системи* починають переписку з іншим користувачем, формуючи діалог, який за допомогою сутностей *Контролер системи* та *База даних* зберігається та існує всередині ІС у вигляді потоку повідомлень конкретної сесії для подальшого динамічного відображення динаміки діалогу даних/повідомлень користувачів/сесій в конкретний момент часу.

16. Сутність *Сервіс системи* надсилає запит сутності *База даних* для перевірки наявності даних, оскільки одним з найважливіших аспектів системи є безпека користувачів та даних, відповідно час від час сутність *Сервіс системи* надсилає сутності *База даних* запит, щоб перевірити цілісність даних та надіслати повідомлення про помилку у разі певних проблем з обробкою даних, також необхідно забезпечити повністю асинхронну обробку та передавання даних заради максимально швидкої роботи ПЗ, що дозволить користувачам оптимально швидко користуватися ІС.

Внутрішня сутність *База даних* складається з основних елементів, які дозволять безпечно та швидко опрацювати інформацію користувача, а саме:

- Перевірити дані.
- Перевірити наявність бекап файлів.
- Зберегти дані.
- Зберегти діалог користувача.
- Сформувати пакет даних.
- Перевірити правильність запиту.
- Відправити дані на запит.

Після цього сутність *Контролер системи* надсилає запит сутності *Сервіс системи* та асинхронну обробку даних. Сутність *Сервіс системи* за допомогою використання динамічних подій постійно здійснює перевірку станів ІС, в яких вона знаходиться і відповідно реагує на всі зміни системи в конкретний момент часу. Сутність *Сервіс системи*

здійснює час від час в окремому потоці перевірку ІС на наявність помилок та винятків, які не були зареєстровані в журналі системи і намагається їх виправити або надіслати на обробку сутності *Контролер системи* на подальше опрацювання. Сутність *Контролер системи* має такі функції:

- Опрацювання даних користувача.
- Опрацювання даних системи.
- Перевірка даних.
- Перевірка доступності сесії.
- Перевірка ідентифікаторів.
- Виведення повідомлення про помилку.
- Зберігання посилань системи.
- Перевірки токена користувача.
- Перевірка правильності оновлень.

Сутність *Сервіс системи* містить такі функції:

- Створення списку користувачів.
- Аналіз списку за допомогою алгоритмів.
- Створення відсоткового співвідношення.
- Обробка переписок користувачів.
- Створення потоку повідомлень.
- Обробка потоку повідомлень.
- Створення користувацьких токенів.

Сутність *Браузер*, хоч і не є основною, але містить такі функції:

- Авторизація/Реєстрація.
- Збереження токена.
- Перевірка сесії.
- Відображення повідомлень користувачу.
- Зчитування користувацьких повідомлень.
- Додавання фотографії користувача.
- Відправлення запитів користувача на сервер.
- Відображення помилок клієнта.

Після того як сутність *Сервіс системи* сформував запит потоку повідомлень, він відправляє запит сутності *Контролер системи*, який запитує токен діалогу в сутності *Сервіс системи*. Остання відправляє поточний токен, в якому містяться дані про поточну сесію. Далі сутність *Контролер системи* через сутність *Браузер* відображає сформований діалог сутності *Користувач*, після чого *Користувач* або продовжує роботу, або виходить з ІС.

Результати досліджень та їх обговорення

Створено діаграму переходів станів, яка відображає систему, що в заданий момент часу знаходиться точно в одному з скінченної множини станів. Діаграма складається з переходів, які містять шість основних етапів:

1. Вхід користувача в систему – користувач входить в ІС, здійснює реєстрацію і/або авторизацію.

2. Додавання фотографії – користувач додає власну фотографію, ІС перевіряє наявність обличчя на фотографії та справжність користувача, здійснює пошук по фотографіях в БД, в разі успішного додавання фотографії зберігає її в БД, а користувач отримує доступ до списків користувачів.

3. Формування списку користувачів – система здійснює аналіз всіх даних користувача та на їх основі здійснює за допомогою алгоритмів нечіткий пошук

по текстовій інформації інших користувачів, тим самим формує список релевантних користувачів по відсотковому співвідношенню, сортує список по спаданню і надсилає на перегляд користувачу.

4. Перегляд вибраного профілю – користувач здійснює перегляд обраного профілю, переглядає фотографії, загальну інформацію та інтереси вибраного користувача і в результаті успіху ставить вподобання профілю та розпочинає переписку.

5. Здійснення переписки – формується окремий потік повідомлень для переписки між двома користувачами, який динамічно оновлюється та зберігається в базі даних системи.

6. Збереження даних та вихід з ІС – після здійснення переписки здійснюється перевірка наявності токена та роботи сесії, зберігаються всі

сформовані дані, користувацька інформація, потік повідомлень та логування в БД системи, після чого користувач виходить в головне меню системи.

На рис. 1 подано композиційну схему ІС, за допомогою якої здійснено поділ об'єктів на класи та відображено залежність класів між собою. На рис. 2 подано діаграму станів, що описує зміни стану об'єкта протягом усього процесу роботи системи. На рис. 3 подано діаграму класифікації видів інформації, а також її підвидів та залежностей необхідних для роботи ІС. На рис. 4 подано діаграму сценарію роботи ІС, яка будує сценарій того, як за допомогою синхронних та асинхронних процесів, змін станів, форм даних та умов виконання запитів функціонує система.

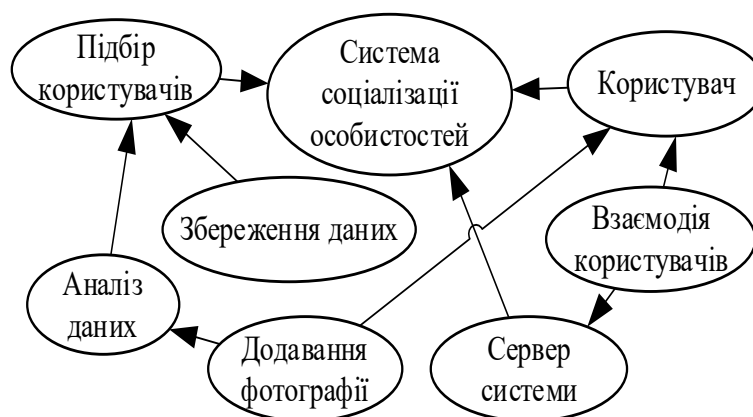


Рис. 1. Композиційна схема інтелектуальної системи

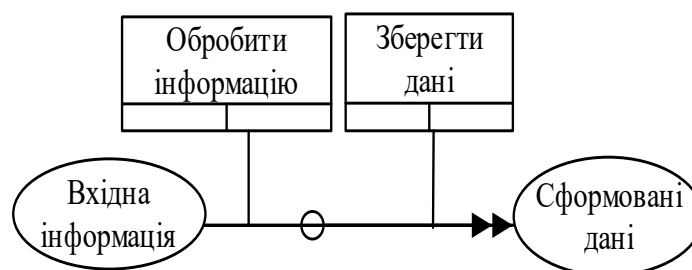


Рис. 2. Діаграма станів



Рис. 3. Діаграма класифікації видів інформації

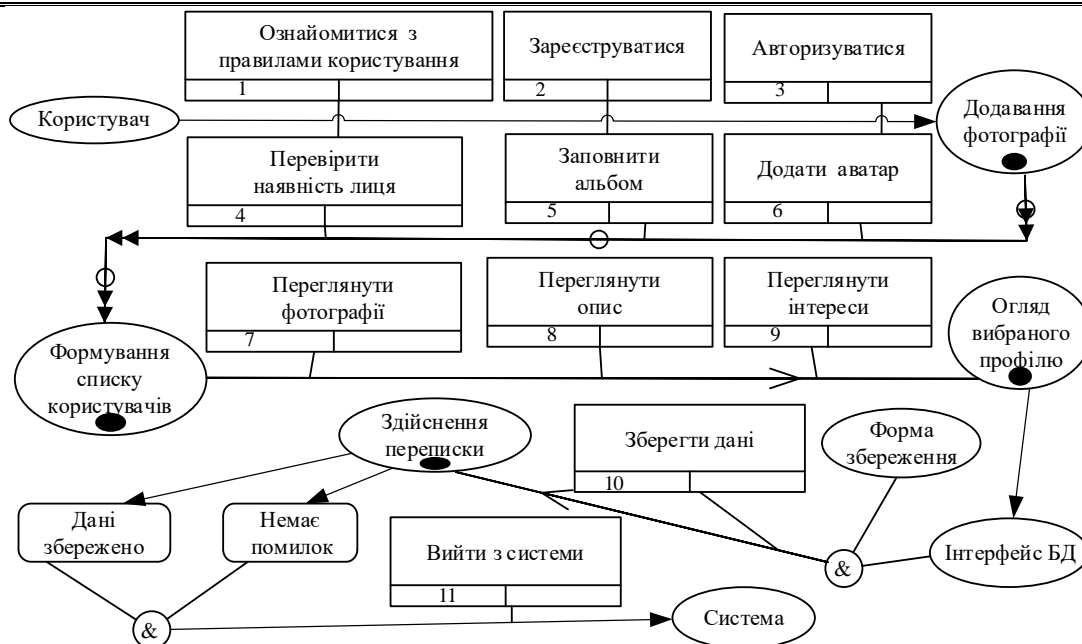


Рис. 4. Діаграма сценарію роботи інтелектуальної системи

Здійснено побудову діаграм сутність-зв'язок, система має наступні сутності:

1. *Користувач* – сутність, що відповідає за додавання особистої інформації, пошук відповідних користувачів ІС, додавання фотографії та здійснення діалогу з вибраним користувачем (рис. 5). *Користувач*

має 5 атрибутів: ID_користувача, Логін, Пароль, Персональні дані, Середній бал. Сутність *Користувач* має три зв'язки з сутністю *Сервер*, а саме: "Шукає користувачів", "Надсилає інформацію" та "Вибирає користувача зі списку".

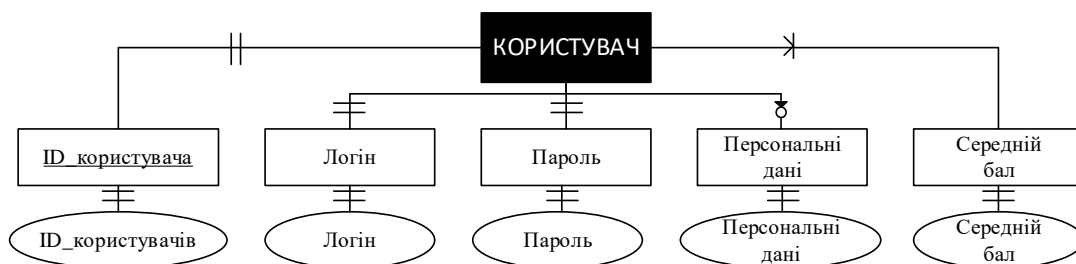


Рис. 5. Сутність "Користувач"

2. *Браузер* – сутність, що відповідає за підтримку інтерактивного діалогу між користувачем та ІС.

3. *База даних* – сутність, що відповідає за збереження даних користувача, підібраних списків релевантних користувачів та проміжних результатів функціонування ІС.

4. *Контролер системи* – сутність, що відповідає за контроль коректного функціонування ІС.

5. *Сервіс системи* – сутність, що відповідає за вибір операцій, здійснення відповідей за запит, виконання алгоритмів аналізу текстової інформації, перевірки справжності користувачів та формування посортowanego списку користувачів по інтересах (рис. 6). Сутність *Сервіс системи* має 5 атрибутів: ID_серверного сеансу, Назва операції, Параметри запиту, Оцінка користувачів, Значення середнього параметру.

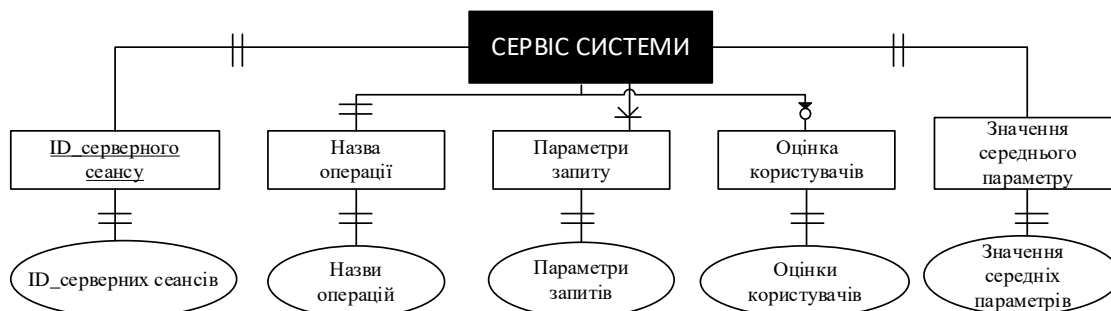


Рис. 6. Сутність "Сервіс системи"

6. *Сесія користувача* – слабка сутність, що відображає сесію поточного користувача та містить всі необхідні параметри для роботи користувача всередині сесії (рис. 7). Сутність *Сесія користувача* є слабкою сутністю, має 2 атрибути: ID_сеансу, Токен

та 2 слабкі зв'язки з сутністю *Користувач*, а саме: "Запит токена" та "Відповідь на запит". *Сесія користувача* має слабкий зв'язок з сутністю *Сервер*, а саме: "Передає дані сесії".



Рис. 7. Сутність "Сесія користувача"

7. *Сервер* – сутність, що сама по собі містить лише сформовані дані та ідентифікатори, але є дуже важливою, оскільки саме завдяки серверу здійснюється взаємодія між іншими сутностями та всі дані в результаті роботи зберігаються на сервері (рис. 8). Сутність *Сервер* має 2 атрибути: Пакет даних, Ідентифікатор.

Сутність *Сервер* має чотири зв'язки з сутністю *Сервіс системи*, а саме: "Зберігає проміжні параметри", "Надсилає запит користувачу", "Зберігає потік діалогу користувачів" та "Формує список користувачів".

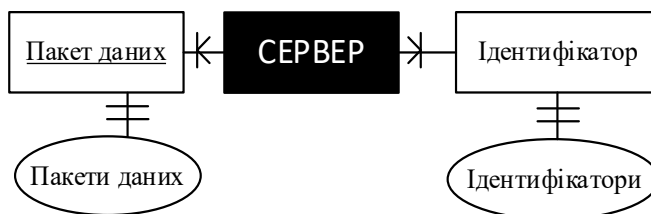


Рис. 8. Сутність "Сервер"

Сутність *Користувач* та слабка сутність *Сесія користувача* мають слабкі зв'язки типу 1-N (рис. 9). Слабка сутність *Сесія користувача* має слабкий зв'язок з сутністю *Сервер* типу 1-N. Сутність

Користувач має зв'язки з сутністю *Сервер* типу 1-N. Сутність *Сервер* має зв'язки з сутністю *Сервіс системи*, а саме: типу 1-N та 1-1.

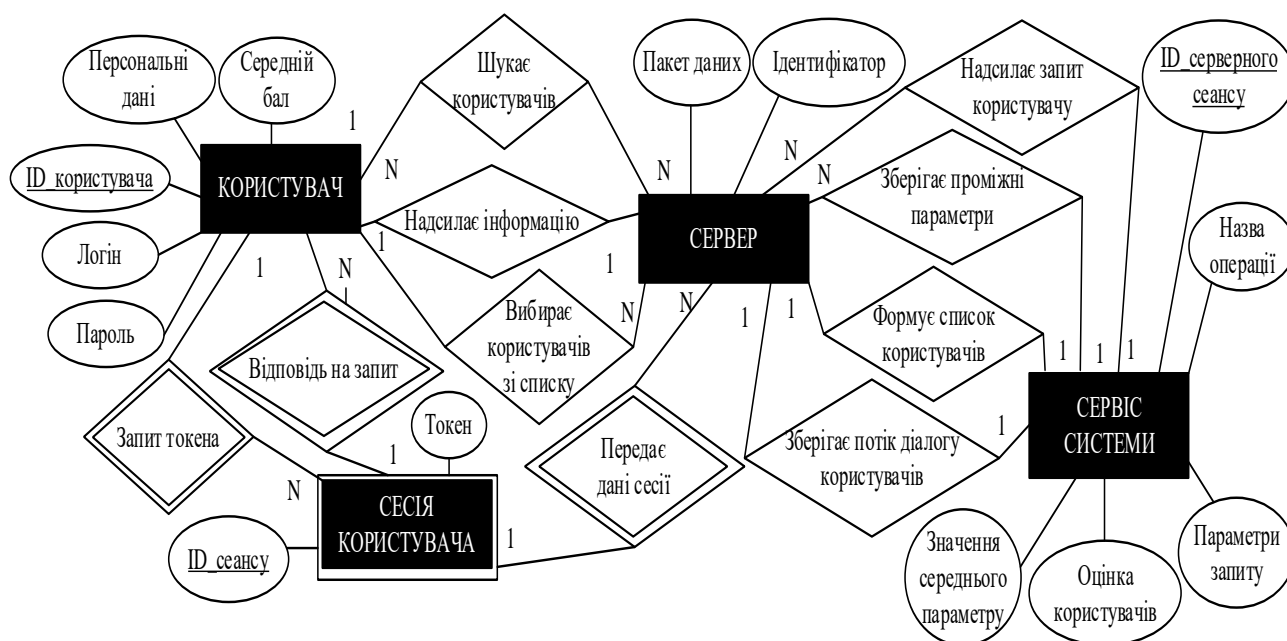


Рис. 9. Діаграма сутність-зв'язок

Створено діаграму переходів станів, яка відображає систему, що в заданий момент часу знаходиться точно в одному з скінченної множини

станів (рис. 10) та розроблено матрицю переходів (табл. 2).

Таблиця 2. Матриця переходів станів системи

Поточний стан	Умова	Дія	Наступний стан
ПОЧАТКОВИЙ СТАН	Вхід користувача в систему	Завантаження системи	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Завантажено систему	Реєстрація	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Зареєстровано	Авторизація користувача	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Авторизовано	Створення токена користувача	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Створено токен	Відправлення токена в браузер на збереження	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Збережено токен	Вхід користувача на головну сторінку	КІНЦЕВИЙ СТАН
ПОЧАТКОВИЙ СТАН	Користувач у системі	Завантаження фотографії	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Завантажено фотографію	Пошук обличчя на фотографії	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Знайдено обличчя	Перевірка обличчя	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Обличчя перевірено	Порівняння обличчя з іншими в БД	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Перевірено обличчя користувача	Збереження фотографії	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Збережено фотографію	Створення альбому	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Додавання фотографії в альбом	Збереження альбому в БД	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Збережено альбом	Завершення роботи з фотографіями	КІНЦЕВИЙ СТАН
ПОЧАТКОВИЙ СТАН	Введено інформацію	Обробка інформації	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Сформовані текстові дані	Обробка тексту	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Оброблено текстові дані	Початок нечіткого пошуку серед всіх користувачів	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Узагальнення нечіткого пошуку між користувачів	Формування списку	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Список сформовано	Сортування списку	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Список посортовано	Фільтрування списку по зовнішніх параметрах	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Відфільтровано список	Вивід інформації користувачу	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Інформація опрацьована	Завершення роботи	КІНЦЕВИЙ СТАН
ПОЧАТКОВИЙ СТАН	Отримано список	Перегляд списку всіх користувачів	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Список переглянуто	Вибір користувача	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Користувача вибрано	Перегляд фотографій користувача	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Фотографії переглянуто	Перегляд опису користувача	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Опис переглянуто	Перегляд інтересів користувача	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Інтереси переглянуто	Виставлення вподобання користувачу	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Вподобання поставлено	Початок переписки з користувачем	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Почато переписку	Створення потоку повідомлень	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Потік створено	Збереження ідентифікатора в БД	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Ідентифікатор збережено	Збереження токена сесії	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Токен сесії збережено	Збереження потоку повідомлень в БД	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Потік збережено	Перевірка токена сесії	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Перевірено наявність токена	Збереження параметрів в БД	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Параметри збережено	Перевірка наявності помилок	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Помилки перевірені	Збереження логування в БД	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Логування збережено	Перевірка локальної БД	ОПРАЦЮВАННЯ
ОПРАЦЮВАННЯ	Локальну БД перевірено	Виведення даних користувачу	ОЧІКУВАННЯ
ОЧІКУВАННЯ	Виведена інформація	Вихід в головне меню	КІНЦЕВИЙ СТАН

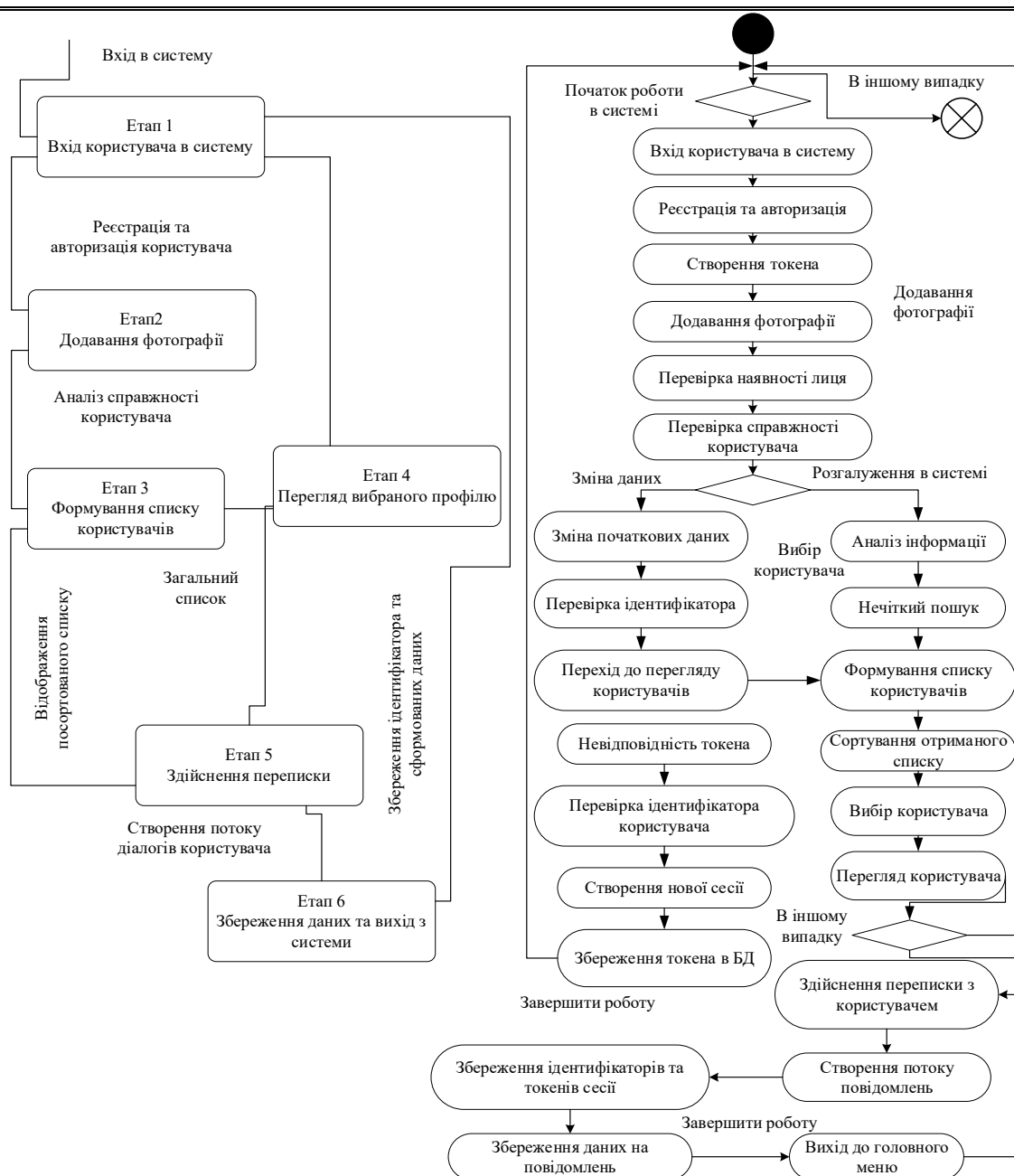


Рис. 10. Діаграма переходів станів

ІС повинна складатися з таких етапів: реєстрація і/або авторизація користувача, знайомство користувача з правилами користування ІС. Завантаження фотографії, яку з використанням згорткової нейромережі перевірено на наявність людського обличчя, та на наявність цього обличчя в вже існуючій БД системи, додавання аватару користувача та заповнення альбому фотографіями. Наступним кроком є вказання користувачем інформації про себе, аналіз вказаної інформації та її порівняння з інформацією про вже існуючих в системі користувачів з використанням алгоритму Левенштейна, алгоритмів нечіткого пошуку та моделі Noisy Channel, після чого формується посортований список користувачів за вирахуванням відсотком схожості між користувачами. Є можливість додатково відфільтрувати сформований список, змінити

особисту інформацію в результаті чого буде наново сформований список, або вийти в головне меню системи. Також всі зміни внесені користувачем, та переписки зберігатимуться в БД, а всі фотографії в хмарному сервісі збереження медіа-даних. Останнім кроком є дослідження профілю та інтересів обраного з сформованої вибірки користувача та початок спілкування з обраним користувачем, вся переписка у вигляді окремого потоку повідомлень зберігається в БД. Загальний алгоритм роботи ІС поданий на рис. 11 у вигляді use case діаграми та на рис. 12 у вигляді activity діаграми. Далі подано основні скріншоти роботи ІС, на рис. 13 подано процес додавання інформації про користувача, а саме опис, інтереси та загальна інформація про те, що та кого шукає користувач.

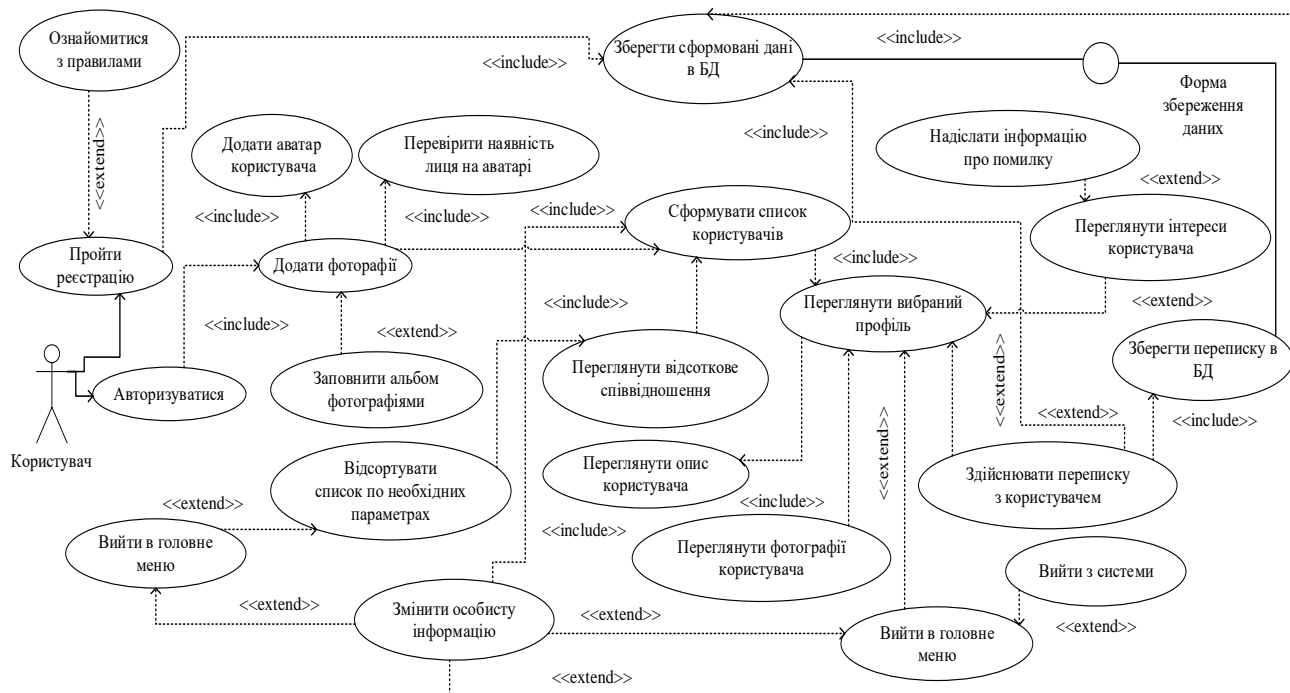


Рис. 11. Use Case діаграма

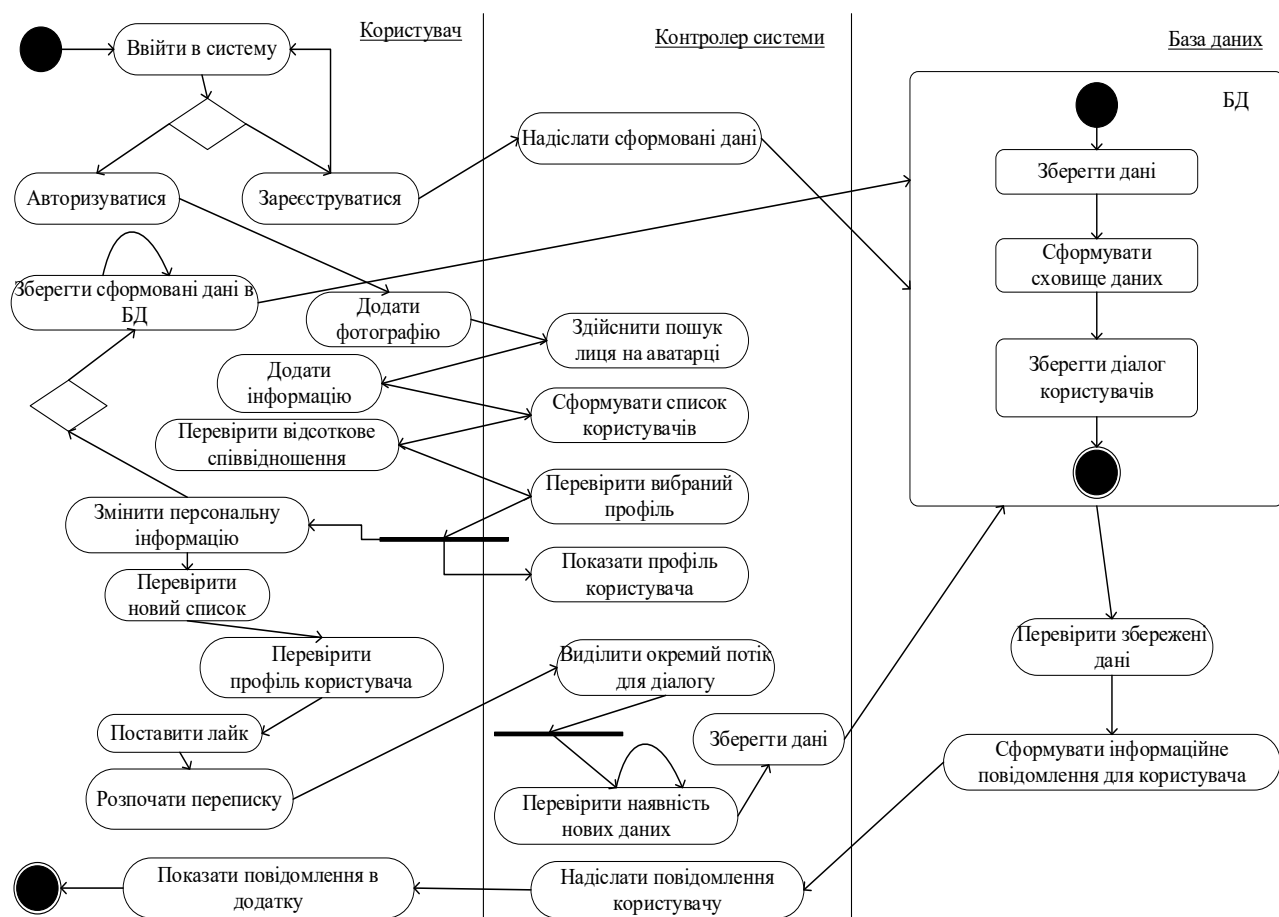


Рис. 12. Activity діаграма

На рис. 14 подано поточні додані фотографії та використання згорткової нейронної мережі. Згорткова нейронна мережа складається з шарів двох типів: згорткових шарів і subsampling шарів або шарів

підвибірки. У ній кожен шар має топографічну структуру, тобто кожен нейрон пов'язаний з фіксованою точкою вихідного зображення, а також з областю вхідного зображення, яка опрацьовується

даними нейроном. У кожному місці кожного шару існує цілий ряд різних нейронів, кожен зі своїм набором вхідних ваг, пов'язаних з нейронами в прямокутному фрагменті попереднього шару. Згортова нейронна мережа дозволяє визначити чи є такий користувач вже в БД, та чи зображене обличчя на фотографії, як видно на виділеному фрагменті скріншоту – 2 картинки система не дозволяє поставити на аватар, оскільки на них немає обличчя.

На рис. 15 подано сформований список користувачів з використанням нечіткого пошуку посортований за вирахуванням відсотком подібності між користувачами. Серед алгоритмів нечіткого пошуку вибрано лінійний пошук – просте послідовне застосування метрики (в даному випадку метрики Левенштейна) до слів з вхідного тексту. Щоб обчислювати оптимальну відстань між словами в

тексті, потрібно модифікувати алгоритм знаходження звичайної відстані Левенштейна наступним чином: зберігати не дві, а три останні рядки матриці, а також додати відповідну додаткову умову – в разі виявлення транспозиції при розрахунку відстані також враховувати її вартість. Таким чином модифікувавши алгоритм використано модель Noisy Channel, що дозволило оптимізувати процес аналізу інформацію, спочатку створивши матрицю параметрів користувачів, потім асинхронно з використанням лінійного пошуку порівнювати поточні параметри і формувати до кожного параметру відсоток подібності. Останнім кроком є пошук середнього відсотку подібності користувачів. На рис. 16-17 подано вибір користувача, перегляд профілю та фотографій обраного користувача та початок переписки.

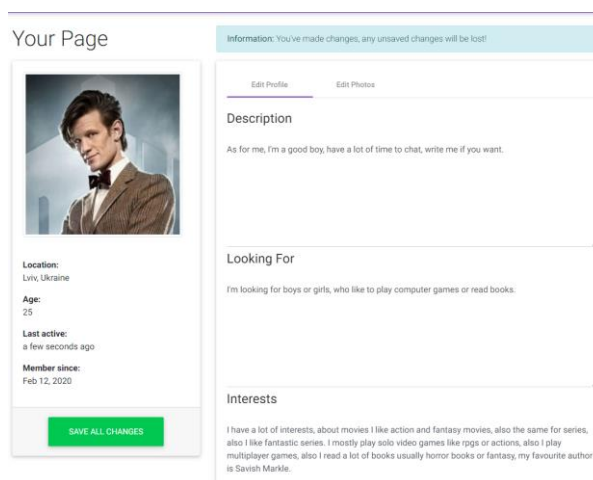


Рис. 13. Інформація користувача

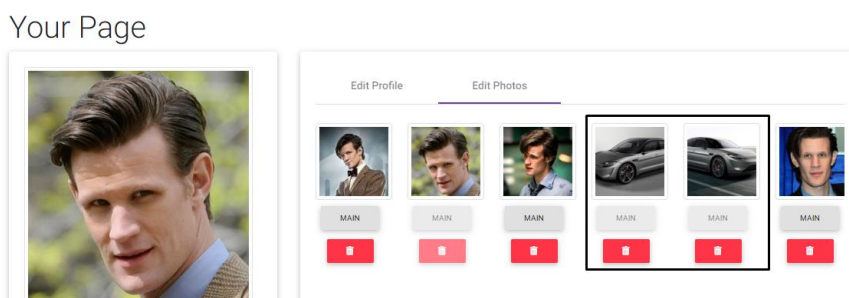


Рис. 14. Вибір аватару користувача

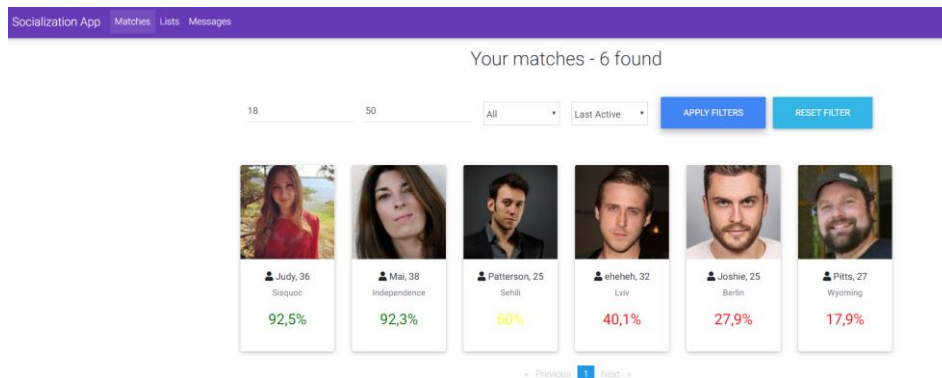


Рис. 15. Сформований список користувачів

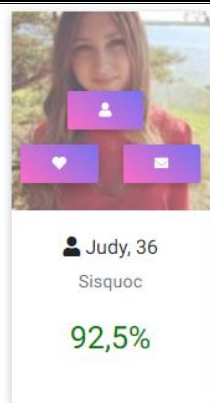


Рис. 16. Вибір користувача

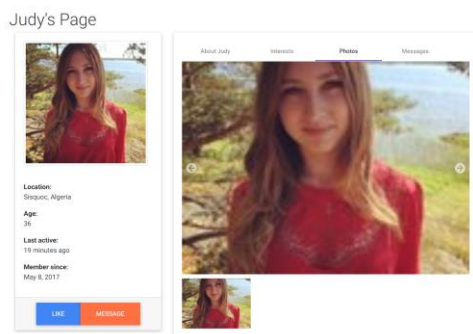


Рис. 17. Перегляд фотографій користувача

Реалізований в ІС алгоритм формування вибірки подібних за інтересами користувачів є ефективнішою та точнішою на 25–30% на відмінну від звичайного алгоритму Левенштейна. Одночасно з цим реалізований алгоритм здійснює вибірку в 10 разів швидше за алгоритм Левенштейна.

Висновки та перспективи подальшого розвитку

Здійснено аналітичний огляд літературних джерел, серед яких коротко описано про всі аспекти

що відносяться до сучасної соціалізації особистостей, а саме інформація про нейронні мережі з розпізнавання обличчя та алгоритми нечіткого пошуку для опрацювання текстової інформації. Описано основні призначення спроектованої та розробленої ІС, мету її створення, які основні проблеми вирішуються при створенні такого типу систем. Проаналізовано причини та фактори, що є важливими для значення актуальності створення даної ІС. Вказані існуючі аналоги системи, їх переваги та недоліки. Здійснено опис основних алгоритмів роботи ІС, а саме алгоритм Левенштейна, модель Noisy Channel та алгоритм N-грам. Наведено приклад роботи алгоритму Левенштейна з пошуком відстані наявної різниці між двома словами. Здійснено побудову необхідних IDEF5 діаграм, в яких описано основні класи системи як об'єкти, створені на основі цих класів та процеси взаємодії об'єктів між собою. Розроблено діаграму переходів станів та відповідно до неї побудовано матрицю переходів станів ІС, де відображено умови, дії та залежності станів системи. Створено діаграму сутність-зв'язок, на якій відображені окремі сутності системи та їх взаємодія в ході виконання системи. Побудовано діаграму варіантів використання та діаграму активностей системи, що дозволило розглянути систему з боку об'єктно-орієнтованого підходу та повноцінно описати загальні вимоги до типової структури створюваної ІС соціалізації особистостей за спільними інтересами, що дає змогу здійснювати подальшу реалізацію відповідного ПЗ. Експериментальна апробація реалізованої ІС соціалізації особистостей на основі розробленого алгоритму формування вибірки спільних подібних за інтересами користувачів продемонструвала точність пошуку релевантних користувачів на 25–30% та швидкість в 10 разів більша при отриманні результату на відмінну від звичайного алгоритму Левенштейна.

Список літератури

1. Parry M. E., Kawakami T., Kishiya K. The effect of personal and virtual word-of-mouth on technology acceptance. *Journal of Product Innovation Management*. 2012. № 29 (6). P. 952–966. DOI: <http://doi.org/10.1111/j.1540-5885.2012.00972.x>
2. Schivinski B., Dąbrowski D. The effect of social-media communication on consumer perceptions of brands. *Journal of Marketing Communications*. 2014. № 22 (2). P. 189–214. DOI: <http://doi.org/10.1080/13527266.2013.871323>
3. Ranjbaran B., Jamshidian M., Dehghan Z. A survey of identification of major factors influencing customers attitude toward machine made carpet brands. *Journal of Business Strategies*. 2007. № 5 (23). P. 109–118.
4. Schmäh M., Wilke T., Rossmann A. Electronic word of mouth: A systematic literature analysis. *Digital Enterprise Computing*. 2017. P. 147–158.
5. Wang X., Yu C., Wei Y. Social media peer communication and impacts on purchase intentions: A consumer socialization framework. *Journal of Interactive Marketing*. 2012. № 26 (4). P. 198–208. DOI: <http://doi.org/10.1016/j.intmar.2011.11.004>
6. Vysotska V. Information Technology for Internet Resources Promotion in Search Systems Based on Content Analysis of Web-Page Keywords. *Radio Electronics, Computer Science, Control*. 2021. № 3. P. 133–151.
7. Batiuk T., Vysotska V., Lytvyn V. Intelligent System for Socialization by Personal Interests on the Basis of SEO-Technologies and Methods of Machine Learning. *Computational Linguistics and Intelligent Systems (COLINS 2020) : 4th International Conference, Lviv, 23-24 April 2020 : CEUR workshop proceedings*. 2020. № 2604. P. 1237–1250.
8. Hudson S., Roth M., Madden T. J. The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism Management*. 2015. № 2 (8). P. 68–76. DOI: <http://doi.org/10.1016/j.tourman.2014.09.001>
9. Hanna R., Rohm A., Crittenden V. L. We're all connected: The power of the social media ecosystem. *Business Horizons*. 2011. № 54 (3). P. 265–273. DOI: <http://doi.org/10.1016/j.bushor.2011.01.007>
10. Guidry J. D., Messner M., Jin Y. From McDonalds fail to Dominos sucks: An analysis of Instagram images about the 10 largest fast food companies. *Corporate Communications: An International Journal*. 2015. № 20 (3). P. 344–359.
11. Gao L. Online consumer behavior and its relationship to website atmospheric induced flow: Insights into online travel agencies in China. *Journal of Retailing and Consumer Services*. 2014. № 21 (4). P. 653–655.

12. Ferrara E., Interdonato R., Tagarelli A. Online popularity and topical interests through the lens of Instagram. *Hypertext and Social Media*. 2014. № 2. P. 24–23. DOI: <http://doi.org/10.1145/2631775.2631808>
13. Erkan I. The influence of e-WOM in social media on consumers' purchase intentions: An extended approach to information adoption. *Computers in Human Behavior*. 2016. № 4. P. 47–55.
14. Elaheebocus S. M., Weal M., Morrison L. Peer-based social media features in behavior change interventions: Systematic review. *Journal of Medical Internet Research*. 2018. № 20 (2). P. 1–20. DOI: <http://doi.org/10.2196/jmir.8342>
15. De-Gregorio F., Sung Y. Understanding attitudes toward and behaviors in response to product placement. *Journal of Advertising*. 2010. № 39 (1). P. 83–96. DOI: <http://doi.org/10.2753/JOA0091-3367390106>
16. Geurin-Eagleman A. N. Communicating via photographs: A gendered analysis of Olympic athletes' visual self -presentation on Instagram. *Sport Management Review*. 2015. № 19 (2). P. 133–145. DOI: <http://doi.org/10.1016/j.smr.2015.03.002>
17. Chu S. C. Using a consumer socialization framework to understand electronic word-of-mouth (eWOM) group membership among brand followers on Twitter. *Electronic Commerce Research and Applications*. 2016. № 14 (4). P. 251–260.

References

1. Parry, M. E., Kawakami, T. (2012), Kishiya, K. "The effect of personal and virtual word-of-mouth on technology acceptance", *Journal of Product Innovation Management*, No. 29 (6), P. 952–966. DOI: <http://doi.org/10.1111/j.1540-5885.2012.00972.x>
2. Schivinski, B., Dąbrowski, D. (2014), "The effect of social-media communication on consumer perceptions of brands", *Journal of Marketing Communications*, No. 22 (2), P. 189–214. DOI: <http://doi.org/10.1080/13527266.2013.871323>
3. Ranjbaran, B., Jamshidian, M., Dehghan, Z. (2007), "A survey of identification of major factors influencing customers attitude toward machine made carpet brands", *Journal of Business Strategies*, No. 5 (23), P. 109–118.
4. Schmäh, M., Wilke, T., Rossmann, A. (2017), "Electronic word of mouth: A systematic literature analysis", *Digital Enterprise Computing*, P. 147–158.
5. Wang, X., Yu, C., Wei, Y. (2012), "Social media peer communication and impacts on purchase intentions: A consumer socialization framework", *Journal of Interactive Marketing*, No. 26 (4), P. 198–208. DOI: <http://doi.org/10.1016/j.intmar.2011.11.004>
6. Vysotska, V. (2021), "Information Technology for Internet Resources Promotion in Search Systems Based on Content Analysis of Web-Page Keywords", *Radio Electronics, Computer Science, Control*, No. 3, P. 133–151.
7. Batiuk, T., Vysotska, V., Lytvyn, V. (2020), "Intelligent System for Socialization by Personal Interests on the Basis of SEO-Technologies and Methods of Machine Learning", *Computational Linguistics and Intelligent Systems (COLINS 2020): 4th International Conference, Lviv, 23-24 April 2020 : CEUR workshop proceedings*, No. 2604, P. 1237–1250.
8. Hudson, S., Roth, M., Madden, T. J. (2015), "The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees", *Tourism Management*, No. 2 (8), P. 68–76. DOI: <http://doi.org/10.1016/j.tourman.2014.09.001>
9. Hanna, R., Rohm, A., Crittenden, V. L. (2011), "We're all connected: The power of the social media ecosystem", *Business Horizons*, No. 54 (3), P. 265–273. DOI: <http://doi.org/10.1016/j.bushor.2011.01.007>
10. Guidry, J. D., Messner, M., Jin, Y. (2015), "From McDonalds fail to Dominos sucks: An analysis of Instagram images about the 10 largest fast food companies", *Corporate Communications: An International Journal*, No. 20 (3), P. 344–359.
11. Gao, L. (2014), "Online consumer behavior and its relationship to website atmospheric induced flow: Insights into online travel agencies in China", *Journal of Retailing and Consumer Services*, No. 21 (4), P. 653–655.
12. Ferrara E., Interdonato R., Tagarelli A. (2014), "Online popularity and topical interests through the lens of Instagram", *Hypertext and Social Media*, No. 2, P. 24–23. DOI: <http://doi.org/10.1145/2631775.2631808>
13. Erkan, I. (2016), "The influence of e-WOM in social media on consumers' purchase intentions: An extended approach to information adoption", *Computers in Human Behavior*, No. 4, P. 47–55.
14. Elaheebocus, S. M., Weal, M., Morrison, L. (2018), "Peer-based social media features in behavior change interventions: Systematic review", *Journal of Medical Internet Research*, No. 20 (2), P. 1–20. DOI: <http://doi.org/10.2196/jmir.8342>
15. De-Gregorio, F., Sung, Y. (2010), "Understanding attitudes toward and behaviors in response to product placement", *Journal of Advertising*, No. 39 (1), P. 83–96. DOI: <http://doi.org/10.2753/JOA0091-3367390106>
16. Geurin-Eagleman, A. N. (2015), "Communicating via photographs: A gendered analysis of Olympic athletes' visual self -presentation on Instagram", *Sport Management Review*, No. 19 (2), P. 133–145. DOI: <http://doi.org/10.1016/j.smr.2015.03.002>
17. Chu, S. C. (2016), "Using a consumer socialization framework to understand electronic word-of-mouth (eWOM) group membership among brand followers on Twitter", *Electronic Commerce Research and Applications*, No. 14 (4), P. 251–260.

Надійшла (Received) 25.12.2021

Відомості про авторів / Сведения об авторах / About the Authors

Батюк Тарас Миронович – магістр кафедри "Інформаційні системи та мережі", Національний університет "Львівська політехніка", вул. Степана Бандери 12, Львів, Україна; e-mail: taras.batiuk.mnsa.2020@lpnu.ua; ORCID ID: <https://orcid.org/0000-0001-5797-594X>

Батюк Тарас Миронович – магістр кафедри "Информационные системы и сети", Национальный университет "Львовская политехника", Украина.

Taras Batiuk – PhD student of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine.

Висоцька Вікторія Анатоліївна – кандидат технічних наук, доцент, доцент кафедри "Інформаційні системи та мережі", Національний університет "Львівська політехніка", вул. Степана Бандери 12, Львів, Україна; e-mail: Victoria.A.Vysotska@lpnu.ua; ORCID ID: <https://orcid.org/0000-0001-6417-3689>

Висоцкая Виктория Анатольевна – к.т.н., доцент, доцент кафедри «Информационные системы и сети», Национальный университет "Львовская политехника", Украина.

Victoria Vysotska – PhD, Associate Professor of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine.

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПОДДЕРЖКИ СОЦИАЛИЗАЦИИ ПОЛЬЗОВАТЕЛЯ ПО СХОЖЕСТИ ИНТЕРЕСОВ

Объектом исследования является процесс социализации личностей, поскольку в наше время задача социализации очень важна и все современные социальные сети пытаются максимально оптимизировать и автоматизировать социализацию разнообразных пользователей с использованием всех популярных современных технологий, таких как нейронные сети и алгоритмы анализа пользовательского текста. **Предметом** исследования является методы и технологические средства поиска и формирования списка релевантных пользователей по сходству интересов для осуществления социализации. Соответственно, осуществляется анализ профиля пользователя системы, а именно определение подлинности пользователя, посредством поиска человеческого лица на пользовательских фотографиях с использованием нейронных сетей и анализ пользовательской информации с использованием алгоритмов нечеткого поиска и модели Noisy Channel. **Цель** работы – создание интеллектуальной системы для социализации личностей на основе нечеткого поиска по словам с использованием модели Noisy Channel с алгоритмами эффективного распределения текстовой информации и сверточной нейронной сети для идентификации пользователей системы. В статье решаются следующие **задачи**: 1. Проанализировать современные и известные подходы, методы, средства и алгоритмы решения задач социализации личностей по сходству интересов. 2. Разработка общей структуры типовой интеллектуальной системы социализации личностей по общим интересам. 3. Сформировать функциональные требования к основным модулям структуры типовой интеллектуальной системы социализации личностей по общим интересам. 4. Разработать интеллектуальную систему поддержки социализации пользователя по сходству интересов на основе нейронных сетей, нечеткого поиска и модели Noisy Channel и провести экспериментальную апробацию. Используются следующие **методы**: метод Левенштейна; модель Noisy Channel; алгоритм N-грам; нечеткий поиск. Получены следующие **результаты**: построена и описана общая структура типовой интеллектуальной системы социализации личностей по общим интересам. К основному назначению создания системы в первую очередь следует отнести создание нового алгоритма анализа пользовательской информации и поиска наиболее подходящих пользователей, в соответствии с проанализированным текстом на основе уже существующих алгоритмов, таких как алгоритм Левенштейна, алгоритм расширения выборки, алгоритм N-грамм и модели Noisy Channel. Также испытан дальнейшее развитие шаблон асинхронного создания программного продукта, что позволит создать практически полностью динамическую систему. Следует усовершенствовать сверточную нейронную сеть, что позволит эффективно и динамично осуществлять поиск человеческих лиц на фотографии и проверять наличие уже существующих людей в базе данных системы. **Выводы**: Выявлено, что реализуемый алгоритм осуществляет выборку примерно в 10 раз быстрее, чем обычный алгоритм Левенштейна. Также реализуемый в системе алгоритм для формирования выборки пользователей является более эффективным и точным примерно на 25-30% по сравнению с обычным алгоритмом Левенштейна.

Ключевые слова: расстояние Левенштейна; модель Noisy Channel; алгоритм N-грам; нечеткий поиск.

INTELLECTUAL SYSTEM DEVELOPMENT FOR USER SOCIALIZATION SUPPORT BY INTERESTS SIMILARITY

The **object** of research is the process of socialization of individuals, because nowadays the task of socialization is very important and all modern social networks try to optimize and automate the socialization of various users using all popular modern technologies such as neural networks and user text analysis algorithms. The **subject** matter of the study is the methods and technologies for the search and formation of a list of relevant users by similarity of interests for socialization. Accordingly, the system user profile analysis is studied, namely the identification of the user by searching the human face in user photos using neural networks and analysing user information using fuzzy search algorithms and the Noisy Channel model. The **goal** of the work is to create an intelligent system for socialization of individuals based on fuzzy word search using the Noisy Channel model with algorithms for efficient distribution of textual information, and a convolutional neural network to identify users of the system. The following **tasks** were solved in the article: 1. Analyse modern and most well-known approaches, methods, tools and algorithms for solving problems of socialization of individuals by similar interests. 2. To development the general structure of a typical intellectual system of socialization of individuals by common interests. 3. To form functional requirements to the basic modules of structure of typical intellectual system of socialization of persons on common interests. 4. Develop an intelligent system of support for user socialization by similarity of interests based on neural networks, fuzzy search and Noisy Channel model and conduct experimental testing. The following **methods** are used: Levenstein's method; Noisy Channel model; N-gram algorithm; fuzzy search. The following **results** were obtained: the general structure of a typical intellectual system of socialization of individuals by common interests was built and described. The main purpose of the system is to create a new algorithm for analysing user information and finding the most suitable users, according to the analysed text based on existing algorithms such as Levenstein's algorithm, sampling algorithm, N-gram algorithm and Noisy Channel model. The template of asynchronous creation of a software product which will allow to create almost completely dynamic system also underwent further development. It is necessary to improve the convolutional neural network, which will allow efficient and dynamic search of human faces in the photo, and check the presence of existing people in the database of the system. **Conclusions**: It was found that the implemented algorithm performs sampling approximately 10 times faster than the usual Levenstein algorithm. Also, the implemented in the system algorithm for forming a sample of users is more efficient and accurate by about 25-30% compared to the usual Levenstein algorithm.

Keywords: Levenshtein distance; Noisy Channel Model; N-gram algorithm; Fuzzy search.

Бібліографічні описи / Bibliographic descriptions

Батюк Т. М., Висоцька В. А. Розробка інтелектуальної системи підтримки соціалізації користувача за подібністю інтересів. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 13–26. DOI: <https://doi.org/10.30837/ITSSI.2022.19.013>

Batiuk, T., Vysotska, V. (2022), "Intellectual system development for user socialization support by interests similarity", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 13–26. DOI: <https://doi.org/10.30837/ITSSI.2022.19.013>

S. MINUKHIN

PERFORMANCE STUDY OF THE DTU MODEL FOR RELATIONAL DATABASES ON THE AZURE PLATFORM

When solving problems of working with relational databases on cloud platforms, the problem arises of choosing a specific model to ensure the performance of executing queries of varying complexity. The **object** of research is the processes of implementing various types of queries to relational databases within the framework of the DTU **purchase** model of the MS Azure platform. The **subject** is methods for evaluating the performance of work with relational databases based on the timing of query execution and indicators of the load on the resources of the cloud platform. The **aim** of the study is to develop a system of indicators for monitoring the current state of work with the database for reasonable decision-making on the choice of a certain price category of the DTU model of the MS Azure cloud service, which will optimize the results of working with the database. Achieving the set goals involves the following **tasks**: to analyze modern tools and services for working with databases, in particular relational databases, on Azure and AWS cloud platforms, the features of their application and implementation; develop software for generating test relational databases of different sizes; test the generated databases on a local resource; taking into account the characteristics of the levels of the Azure DTU model, develop a new system of performance indicators, which includes 2 groups - time indicators and indicators of the load on existing platform resources; develop and implement queries of varying complexity for the generated test database for different levels of the DTU model and analyze the results. **Methods.** The following methods were used in the research: methods of relational database design; methods of creating queries in SQL-oriented databases with any number of tables; methods of creating and migrating data to cloud platforms; methods of monitoring the results of queries based on time and resource indicators; methods of generating test data for relational databases; system approach for complex assessment and analysis of productivity of work with relational databases. **Results.** On the basis of the developed scorecard used for the current analysis of the processes of working with relational databases of the MS Azure platform, numerous experiments were carried out for different levels of the model for simple and complex queries to a database with a total volume of 20 GB: loading of DTU model levels when executing various queries, the influence of model levels DTU Azure SQL database on the performance of simple and complex queries, the dependence of the execution time of various queries on the load of the CPU and the speed of write/read operations for different levels of the model. **Conclusions.** The results of the experiments allow us to conclude that the levels of the DTU model - S3 and S7 - are used to generate test data of various sizes (up to 20 GB) and execute database queries. The practical use of the proposed indicators to evaluate the results of applying the DTU model will improve the efficiency of decision-making on choosing the model level when implementing various queries and generating test data on the Azure cloud platform. The developed set of indicators for working with relational databases on the Azure cloud platform expands the basis of the methodological framework for evaluating the performance of working with relational databases on cloud platforms by analyzing the results of executing the simple and complex database queries on the resources involved.

Keywords: cloud platform; relational database; DTU purchase model; indicators of time and workload; data generator; test data; request complexity.

Relevance

One of the most powerful and relevant trends in the use of modern information and communication systems of different spheres of application is the creation and implementation of database technologies. The use of these technologies allows a significant increase in the efficiency of information systems such as the Internet of Things, cloud platform services, distributed data processing systems – in distributed computing systems (e.g., grid systems for intensive data processing), clusters with the ability to connect data processing centers (DPCs) using service-oriented archetypes, etc. One of the tasks here is to create and test relational databases based on different tools. Relational databases are now not only one of the most popular technologies in terms of their versatility in solving a significant number of tasks of various purposes, but also a sufficiently large variety of their proposals in the market of information services. DB development is an important stage of its use – the creation of logical and physical models is accompanied by constant monitoring of the status of the existing database for optimization of their work in conditions of scale and dynamic changes in the number of users. This raises the problem of locating and storing the database on remote servers, providing access to them, which requires compliance with the requirements for communication systems (the capabilities of Internet

resources), data center capacity, as well as the associated financial costs. Therefore, there is a need for a general solution to this problem and related partial tasks. One of the possible ways is to use the services of the data storage platforms designed to work with databases and data warehouses

A brief overview of cloud platform services

Cloud computing is one of the trends of modern information technology, which is rapidly developing and is increasingly used at different levels of management. Cloud computing is the provision of computing capacities, warehouses, databases, resources and other resources of cloud service platforms via the Internet [1]. Cloud computing has the following advantages [2]:

1. Capital expenditures are being converted into changes. There is no need to invest large sums of money in data processing centers (DPC) and servers, not knowing in advance what capacities will be needed in the future. You can pay only for the actual use of computing resources.

2. Significant savings in large volumes. When using cloud computing, you can achieve a lower variable cost than when creating your own computing infrastructure.

3. There is no need to predict what amount of infrastructure resources will be needed. When making a

decision on the volume of resources for program implementation, often in the final time you can run into downtime of expensive resources or lack of capacity.

4. Increasing the speed and flexibility of developing the features.

5. Lack of costs for launching and maintaining DPC resources.

The global market for cloud solutions and services is growing quite rapidly, so it is difficult to predict the rate of its increase in practice. However, the existing data record the same trends: the rapid growth rate of cloud computing costs, as well as the associated market for services, data centers and data traffic in these systems [1, 2]. In April 2018, Gartner analysts [3] published the results of a study of the global public cloud market. The

cost of these solutions is increased (table 1) due to infrastructure services (IaaS model – DigitalOcean, Linode, Rackspace, Amazon Web Services (AWS), Cisco Metapod, Microsoft Azure, Google Compute Engine (GCE).) And software provided as a service (SaaS model - AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, OpenShift) [1, 2, 4]. In 2021, the cost of cloud services is projected to increase almost 2 times compared to 2017. The largest segment of the public cloud market remains SaaS, which, according to Gartner, by 2021 will account for 45% of software costs in the world. The fastest growing segment of analytics is the IaaS service model, the volume of which in 2018 increased by almost 40%.

Table 1. World forecast of cloud services (billion dollars)

Cloud services	2017	2018	2019	2020	2021
Business Process as a Service (BPaaS)	42,6	46,4	50,1	54,1	58,4
Platform as a Service (PaaS)	11,9	15,0	18,6	22,7	27,3
Software as a Service (SaaS)	60,2	73,6	87,2	101,9	117,1
Cloud management and security services	8,7	10,5	12,3	14,1	16,1
Infrastructure as a Service (IaaS)	30,0	40,8	52,9	67,4	83,5
Total market	153,5	186,4	221,1	260,2	302,5

In recent years, cloud services such as DBaaS (Database as a Service, AWS) and MWaaS (Middleware as a Service) have appeared, which are a type of PaaS. Using DBaaS, the user can access any type of database on request and quickly deploy databases on any hardware in the environment of the selected software platform (operating system). MWaaS is a specialized cloud solution for companies. This service provides access to a comprehensive platform with the appropriate infrastructure to service enterprise programs and security tools. Using MWaaS, the user can quickly prepare a specific software environment to perform the appropriate tasks.

The highest rates of DBaaS use are based on the Azure SQL database [5, 6, 9] and RDS AWS [7, 8]. Tools and technologies related to the use of the following models have been developed: deployment models, migration models, and purchase models. At present, these models are in fact the stages of creation, migration and use of a certain price category of models used for relational DB on cloud platforms [10-12].

These models of services for working with relational DB on cloud platforms to justify the choice of resources when working with certain DB do not fully take into account the nature of work with DB, namely, the volume of DB (number of records in tables), its filling method (location) migration from a local resource or direct deployment to available Azure resources), DB complexity (number and volume of tables), DB query types - simple and complex queries – that significantly affect the time and efficiency of working with DB. This is especially important when scaling DB and when using so-called

"cold" and operational data in the process of analytical data processing of enterprises and institutions. Therefore, it is key and promising to develop indicators for quantitative (qualitative) evaluation of the results of various types of queries to DB, taking into account practical issues related to the scalability of systems (data volumes) based on existing price categories of cloud platform models.

Analysis of publications and setting the task

Let's analyze the state of the problem on the basis of DB servers on the Azure and AWS cloud platforms [5 – 8].

Microsoft Azure cloud platform: characteristics and models of work with relational DB [5, 6].

Azure database SQL resides in the Azure cloud and is part of the platform-as-a-service (PaaS) server model. The SQL Azure database allows you to easily purchase a fully managed DBMS PaaS core that meets your productivity and cost requirements. Depending on the model of SQL Azure data base deployment, one can choose the model that meets the user's needs [5, 6]:

1. Based on DTU units (Database Transaction Units). This model offers logical servers in SQL Azure database.

2. On the basis of virtual cores (vCore) – this model offers logical servers in SQL Azure database and kernal instances in SQL Azure database.

Table 2. Comparison of types of purchase models on the Microsoft Azure cloud platform

Purchase model	Description	Optimal
Purchase model based on DTU units	It is based on complex evaluation of computing resources, storage resources and resources for input-output operations. The productivity levels are expressed in data base transaction units (DTU) for individual databases, and for elastic pools - in elastic data base transaction units (eDTU).	For customers who need simple, pre-configured resource options.
Purchase model based on virtual cores	This model allows to independently choose computing resources and storages. In addition, it allows you to save costs due to the "Advantage of Azure Hybrid Usage for SQL Server" program.	For customers who value flexibility, control and transparency.

The DTU-based purchase model offers a number of pre-configured computing resource packages and included storage capacity to ensure different levels of productivity of add-ons. Data base transaction unit (DTU) is an integral indicator of the capacity (productivity) of the CPU, RAM, read and write operations and memory on a hard disk drive (HDD, SSD). The physical characteristics (CPU, memory, input/output operations) associated with each

DTU measurement are calibrated using the performance test which has the operating load of the real data base. The productivity test consists of a set of transactions of different types, which are executed in a data model containing a number of tables and data types. When using a model based on DTU units, users can choose between different service levels (Basic, Standard and Premium) for certain databases and elastic pools [5] (tables 3, 4).

Table 3. Levels of DTU-based purchase models

Indicator	Basic	Standard	Premium
Target workload	Development and application in the working environment.	Development and application in the working environment.	Development and application in the working environment.
Service level agreement with a guarantee of continuous operation time	99.99%	99.99%	99.99%
Backup storage	7 days	35 days	35 days
CPU	Low	Low, medium, high	Medium, high
I / O bandwidth (approx.)	2.5 I/O operations on the DTU	2.5 I/O operations on the DTU	48 I/O operations on the DTU
I / O delay (approximately)	5 ms (read), 10 ms (write)	5 ms (read), 10 ms (write)	2 ms (read and write)
Indexing columnstore	Not available	S3 and above	Supported
In-memory OLTP	Not available	Not available	Supported
Maximum storage size	2 GB	1 TB	4 TB
Maximum number of DTUs	5	3,000	4,000

Table 4. Characteristics of levels of the purchase model based on DTU units

Level, model		DTU	Storage volume included, GB	Max. storage volume, GB	Max. OLTP volume in memory, GB	Max. number of queries, that are performed simultaneously, in 1 second.	Max. number of simultaneous sessions
1		2	3	4	5	6	7
Basic	B	5	2	2	N/A	30	300
Standard	S0	10	250	250	N/A	60	600
	S1	20	250	250	N/A	90	900
	S2	50	250	250	N/A	120	1200
	S3	100	250	250, 500, 750, 1024	N/A	200	2400
	S4	200	250	250, 500, 750, 1024	N/A	400	4800
	S6	400	250	250, 500, 750, 1024	N/A	800	9600
	S7	800	250	250, 500, 750, 1024	N/A	1600	19200
	S9	1600	250	250, 500, 750, 1024	N/A	3200	30000
	S12	3000	250	250, 500, 750, 1024	N/A	6000	30000

The virtual core purchase model (vCore) provides flexibility, control, and transparency in the consumption of individual resources. A virtual core is a logical central processor (CPU) that can select hardware generation and physical hardware characteristics (such as number of cores, memory capacity, and storage size). The model allows you to **independently** choose computing and

storage resources, ensure performance at the local environment level and optimize costs and uses two levels of services: general purpose and critical for business [5]. Service levels differ in the range of performance levels, the structure to ensure a high level of availability, the method of providing isolation from failures, types of storage and the range of input- output operations (table 5).

Table 5. Levels of vCore-based purchase models

Characteristics	General purpose	Critically important for business
1	2	3
Optimal for	Most workloads. Offers budget balanced and scalable variants of computing and storage resources.	Business applications with high demands on input-output operations. Offers the highest resistance to failures due to the use of several isolated replicas.
Computing services	- 4th generation: from 1 to 24 virtual cores;	- 4th generation: from 1 to 24 virtual cores;
Memory	- 5th generation: from 1 to 80 virtual cores.	- 5th generation: from 1 to 80 virtual cores.
Input/output bandwidth (speed)	- 4th generation: 7 GB per core;	- 4th generation: 7 GB per core;
Storage service	- 5th generation: 5.1 GB per core.	- 5th generation: 5.1 GB per core
Availability	500 input/output operations per second on the virtual core with a maximum of 7,000 operations per second.	5,000 input/output operations per second per core with a maximum of 200,000 operations per second
Backups	- Premium level remote control storage;	- Local storage on SSD storage devices;
In memory	- individual database: 5 GB - 4 TB;	- separate database: 5 GB - 1 TB;

Amazon Web Services cloud platform: characteristics and models of working with relational DB [7, 8].

Amazon Relational Database Service (Amazon RDS) is a web service that allows you to set up, use, and scale relational databases in Amazon Web Services cloud. Amazon RDS is available in the form of database

instances of several types: optimized for working with memory, for a high productivity or implementation of the write/read operations (table 6) and offers a choice of six database cores, including Microsoft SQL Server. When working with the SQL Server database on the AWS platform, the deployment model, the migration model and the purchase model are also used.

Table 6. SQL Server database instance classes of the Amazon RDS service

Instance type	vCPU processors	Memory, GB	Optimized for IOPS (write/read speed)	Instance type
Standard - the last generation				
db.m4.large	2	8	Yes	Medium
db.m4.xlarge	4	16	Yes	High
db.m4.2xlarge	8	32	Yes	High
db.m4.4xlarge	16	64	Yes	High
db.m4.10xlarge	40	160	Yes	10 gigabits
Standard - the previous generation				
db.m3.medium	1	3,75	–	Medium
db.m3.large	2	7,5	–	Medium
db.m3.xlarge	4	15	Yes	High
db.m3.2xlarge	8	30	Yes	High
Memory optimization - current generation				
db.r3.large	2	15	–	Medium
db.r3.xlarge	4	30,5	Yes	Medium
db.r3.2xlarge	8	61	Yes	High
db.r3.4xlarge	16	122	Yes	High
db.r3.8xlarge	32	244	–	10 gigabits
Microinstances				
db.t2.micro	1	1	–	Low
db.t2.small	1	2	–	Low
db.t2.medium	2	4	–	Medium
db.t2.large	2	8	–	Medium

You do not need to purchase Microsoft SQL Server licenses separately for service models that include a license. The cost of licensed models includes the cost of software, basic hardware resources and Amazon RDS control capabilities. This model supports: Express Edition, Web Edition, Standard Edition and Enterprise Edition [7]. The free use of AWS for Amazon RDS allows you to use micro DB instances for free in one availability zone running SQL Server Express Edition. By default, users can create up to 40 instances of the Amazon RDS database, including up to 10 instances of the SQL Server DB under the licensing model. Therefore, the functionality of using such a service is quite limited from the point of view of the average user of cloud platforms.

Thus, the analysis and literary sources revealed the advantages of the Microsoft Azure platform over the Amazon Web Services platform:

1. Azure gives 200 dollars for 30 days as a free subscription, which allows you to use a wide range of services during this period. This allows us to investigate the impact of increasing the price category of the database on its productivity. AWS does not have such a possibility.

2. Access to the Azure database can be obtained using the built-in editor of records in Azure Portal, Microsoft SQL Server Management Studio, and Microsoft Visual Studio. Access to the SQL Server database, located on the AWS platform, is available only through Microsoft SQL Server Management Studio.

3. With more regions available than any other service provider, the Azure platform offers scale and regions to keep additions as close as possible to users.

4. Microsoft Azure has a more intuitive interface for creating and working with databases.

5. Microsoft Azure provides databases with extensive real-time monitoring capabilities, automated performance tuning based on inverse access models and requests performance analysis.

6. Microsoft Azure has a more flexible interface for working with relational databases and provides the ability to easily set up and smoothly switch from one level (price category) to another in the DTU model (within S0-S12) (see table 3, table 4) as compared to the vCore model.

The purpose of this article is to develop indicators for assessing and analyzing the performance of the relational database service and to investigate the use of price categories of existing models for different types of queries (on the basis of the Azure cloud platform).

Materials and methods

Algorithm of the research conducting.

The algorithm was developed according to the given sequence of stages of work with the database on the cloud platform **with such modification**: designing and deploying a relational database, developing queries of varying complexity to the database, generating test data of slow database volume and conducting experiments for the DTU purchase model at the S0-S12 levels of the SQL database service.

The algorithm is implemented by the following steps.

Step 1 Creation of SQL Server database (projecting).

Step 2 Deployment of SQL Server database on local resource.

Step 3. Filling SQL Server database with test data using random number generator. The generator allows to dynamically change the number of records in the created database tables for evaluation of the influence of their volume on the performance indicators of the model DTU.

Step 4: Development of simple and complex types of queries for further evaluation of their results for some level of the DTU Azure SQL database model.

Step 5: Testing the created database on a local resource.

Step 6. Development of relational database productivity indicators for DTU Azure SQL database model taking into account characteristics of the purchase model levels on the basis of DTU units (see tables 3 and 4).

Step 7: Analysis of the results obtained by performing different types of queries to the database, and formulating recommendations for choosing the price category of the purchase model DTU Azure SQL database for different types of queries.

Database design.

2 tables with a size of 1,000,000 records each (10 GB) were designed for testing (tables 7, 8).

Table 7. Customer

Field	Data type
Customer ID	unique identifier, Primary key
Email	nvarchar(100)
Password	nvarchar(50)
Creation Date	Date time

Table 8. Permission

Field	Data type
Permission ID	unique identifier, Primary key
Customer ID	unique identifier, Foreign key
CanRead	Bit
CanWrite	Bit

Performance indicators for relational databases for the DTU Azure SQL database purchase model.

To monitor the performance of the Azure SQL database and the managed instance of Azure SQL, a system for monitoring CPU usage and CPU usage is used to control a certain level of the database model on an ongoing basis. To do this, the Azure SQL database and the managed instance of Azure SQL in the process can provide resource metrics that can be viewed on the Azure portal or using Azure Data Studio or SQL Server Management Studio (SSMS). However, these indicators are not enough for a comprehensive analysis of the use of certain levels of models of database work on the Azure platform, for which it is proposed to expand their composition taking into account the characteristics of the DTU model.

Based on the analysis of characteristics [9, 10, 13] and indicators of levels (price categories) of the purchase model DTU Azure SQL database [14, 15] the following

indicators are proposed for evaluation and analysis of its performance (table 9).

Table 9. Database performance indicators for the DTU purchase model

№	Characteristic	Description	Comment
1	Request type	Determined by the level of complexity of the query (simple, complex)	Determined by the number of tables and poles in the query
2	Query execution time	Query type execution time at DTU model level, sec	It is determined by the execution time of a certain type of query at the model level
3	The number of selected records in the query	Number of selected records from fixed-volume database tables for simple and complex queries	The result of the query,
4	The ratio of the number of selected records to the total number of records in the tables, %	The effectiveness of the query	absolute
5	DTU load, %	DTU model level load	The result of the query,
6	CPU load, %	CPU usage for DTU model level	relative
7	Load OP, GB	RAM load capacity, GB	Determined by the load level of the DTU model for the query type
8	Number of operations per 1 second	Number of query execution operations per 1 sec. (input / read speed)	The CPU load for the DTU model level for the request type is determined
9	Number of DTUs on request	Must not exceed 80% for all levels of the DTU model	Determined by the amount of OP used to execute the query

The first 4 indicators are time indicators and performance of requests; the rest are indicators of resource utilization in the execution of requests.

Assessing the impact of DTU Azure SQL database model levels on the performance of simple and complex queries to relational databases.

3 price categories (levels) of the model were selected to conduct research to assess the impact of price categories of the DTU purchase model: S0, S3 and S7.

The results of execution of different types of queries for performance indicators of the DTU model "Execution time", "Number of selected records per query" and "Ratio of the number of selected records to their total number" are summarized in tables 10 – 12 and in fig. 1.

Table 10. Results of execution of queries to the database for the price category S0

№ of the query	Description of the query	Table	Execution time, sec.	Number of selected records on the query	The ratio of the number of selected records to their total number, %
Simple queries (SQ)					
SQ 1	SELECT * FROM dbo.Customers	dbo.Customer	79	1001002	50,03
SQ 2	SELECT * FROM dbo.Customers WHERE Email LIKE 'A1%'	dbo.Customer	2	706	0,04
SQ 3	SELECT * FROM dbo.Customers WHERE CreationDate < GETDATE()	dbo.Customer	93	201002	10,05
Complex queries (CQ)					
CQ 1	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId	dbo.Customer, dbo.Permission	105	1001002	50,03
CQ 2	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.Password LIKE 'A%' AND per.CanRead = 1	dbo.Customer, dbo.Permission	54	13827	0,69
CQ 3	SELECT * FROM dbo.Customers AS cus LEFT JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.CreationDate < GETDATE()	dbo.Customer, dbo.Permission	103	201002	10,05

The results of queries for the S3 category of the model are given in table 11.

Table 11. Results of database queries for price category S3

№ of the query	Description of the query	Table	Execution time, sec.	Number of selected records on the query	The ratio of the number of selected records to their total number, %
Simple queries (SQ)					
SQ 1	SELECT * FROM dbo.Customers	dbo.Customer	73	1001002	50,03
SQ 2	SELECT * FROM dbo.Customers WHERE Email LIKE 'A1%'	dbo.Customer	1	706	0,04
SQ 3	SELECT * FROM dbo.Customers WHERE CreationDate < GETDATE()	dbo.Customer	99	201002	10,05
Complex queries (CQ)					
CQ 1	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId	dbo.Customer, dbo.Permission	97	1001002	50,03
CQ 2	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.Password LIKE 'A%' AND per.CanRead = 1	dbo.Customer, dbo.Permission	43	13827	0,69
CQ 3	SELECT * FROM dbo.Customers AS cus LEFT JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.CreationDate < GETDATE()	dbo.Customer, dbo.Permission	54	201002	10,05

The results of queries for the S7 model category are given in table 12.

Table 12. Results of database queries for price category S7

№ of the query	Description of the query	Table	Execution time, sec.	Number of selected records on the query	The ratio of the number of selected records to their total number, %
Simple queries (SQ)					
SQ 1	SELECT * FROM dbo.Customers	dbo.Customer	53	1001002	50,03
SQ 2	SELECT * FROM dbo.Customers WHERE Email LIKE 'A1%'	dbo.Customer	1	706	0,04
SQ 3	SELECT * FROM dbo.Customers WHERE CreationDate < GETDATE()	dbo.Customer	74	201002	10,05
Complex queries (CQ)					
CQ 1	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId	dbo.Customer, dbo.Permission	84	1001002	50,03
CQ 2	SELECT * FROM dbo.Customers AS cus INNER JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.Password LIKE 'A%' AND per.CanRead = 1	dbo.Customer, dbo.Permission	37	13827	0,69
CQ 3	SELECT * FROM dbo.Customers AS cus LEFT JOIN dbo.Permissions AS per ON cus.CustomerId = per.CustomerId WHERE cus.CreationDate < GETDATE()	dbo.Customer, dbo.Permission	42	201002	10,05

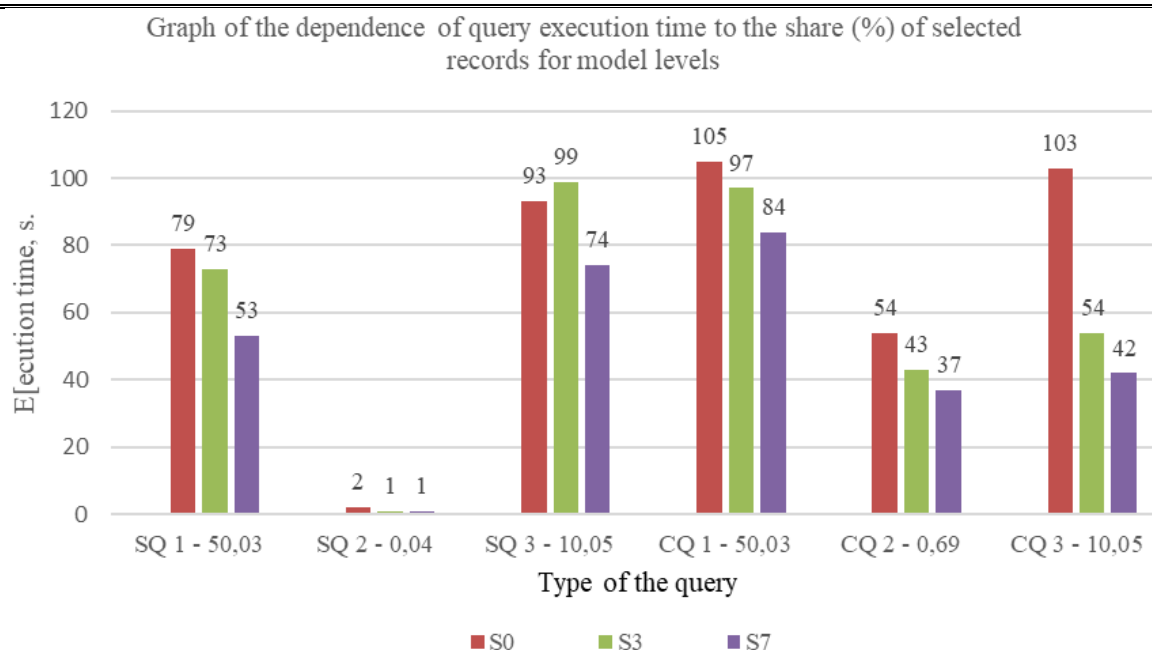


Fig. 1. Graph of query execution time dependence on the share of selected records for different queries, %, for DTU model levels

From the graph we can conclude that the execution time of different types of queries for their performance for different levels of the model varies in the range: for **simple queries** – 1.3 - 1.5 times, for **complex queries** - 1.25 - 2.5 times. Thus, increasing the level of the model allows to increase the productivity of the database in the range of 130% - 250% (for databases of 20 GB), and the complexity of the query is determined by the relative indicator – **the number of received records for queries**

to the total database volume **during the time interval of their implementation.**

Further analysis of the model's performance will be based on the **indicators of resource load**: "DTU load", "CPU load", "RAM load" and "Number of operations per 1 sec." (see table 9).

The load levels of the DTU model when generating database test data are given in table 13.

Table 13. Indicators of the load levels of the DTU model when generating test data of the database

Model level	DTU load, %	CPU load, %
S0	100	15,57
S3	100	3,36
S7	55	0,72

The load levels of the DTU model when performing various queries are given in table 14.

Table 14. Load rate levels of the DTU model when executing different types of queries

Model level	DTU load, %	CPU load, %	RAM load, GB.	Number of operations per 1 sec.
1	2	3	4	5
Query SQ1				
S0	44	4,2	1,17	12671
S3	19	1,82	1,17	13712
S7	1,7	0,08	1,17	18887
Query SQ2				
S0	52	3,8	0,9	353
S3	23	3,85	0,9	706
S7	2	0,31	0,9	706
Query SQ3				
S0	68	5,2	1,05	2161
S3	22	1,54	1,05	2030
S7	1,8	0,9	1,05	2716

The end **Table 14**

1	2	3	4	5
Query CQ1				
S0	94	4,95	1,43	9533
S3	37	1,32	1,43	10320
S7	3,7	0,35	1,43	11917
Query CQ2				
S0	83	5,74	1,22	256
S3	32	3,57	1,22	322
S7	1,4	0,21	1,22	374
Query CQ3				
S0	81	4,3	1,14	1951
S3	31,7	2,69	1,14	3722
S7	2,1	0,31	1,14	4786

Thus, for different types of queries, the value of the key indicator -load of DTU varies: for simple queries – in the range of 1.8% - 68%, complex queries - in the range of 1.4% - 94%, and therefore the optimal level of the S7 model.

Table 15. Number of operations per second (speed) when executing queries for level S0

Query type	Number of selected records	Query execution time, sec.	Number of operations per second
SQ 1	1001002	79	12671
SQ 2	706	2	353
SQ 3	201002	93	2161
CQ 1	1001002	105	9533
CQ 2	13827	54	256
CQ 3	201002	103	1951

Table 16. Number of operations per second (speed) when executing queries for level S3

Query type	Number of selected records	Query execution time, sec.	Number of operations per second
SQ 1	1001002	73	13712
SQ 2	706	1	706
SQ 3	201002	99	2030
CQ 1	1001002	97	10320
CQ 2	13827	43	322
CQ 3	201002	54	3722

Table 17. Number of operations per second (speed) when executing queries for level S7

Query type	Number of selected records	Query execution time, sec.	Number of operations per second
SQ 1	1001002	53	18887
SQ 2	706	1	706
SQ 3	201002	74	2716
CQ 1	1001002	84	11917
CQ 2	13827	37	374
CQ 3	201002	42	4786

Thus, the speed of input/read operations when performing various queries to the database varies in the ranges: for simple queries 12671 (S0) – 18887 (S7), for complex queries – 256 (S0) – 374 (S7). Therefore, the choice of the S7 model level is optimal.

Figures 2 - 4 show the results of visualization to analyze the relationships between query execution time and resource indicators used in the model. They show that it is desirable to use S2 level for simple queries and at least S3 level for complex queries.

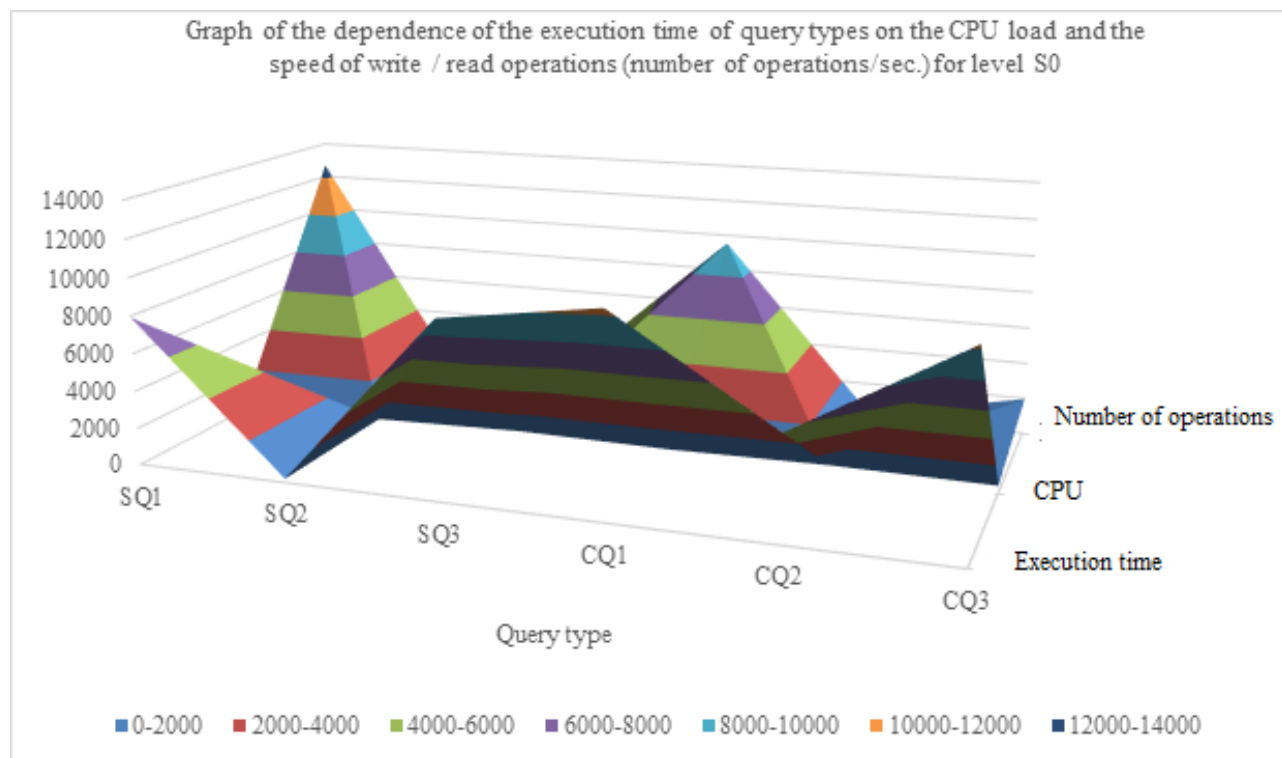


Fig. 2. Graph of the dependence of query execution time on CPU load and write / read operation speed (number of operations/sec.) for level S0

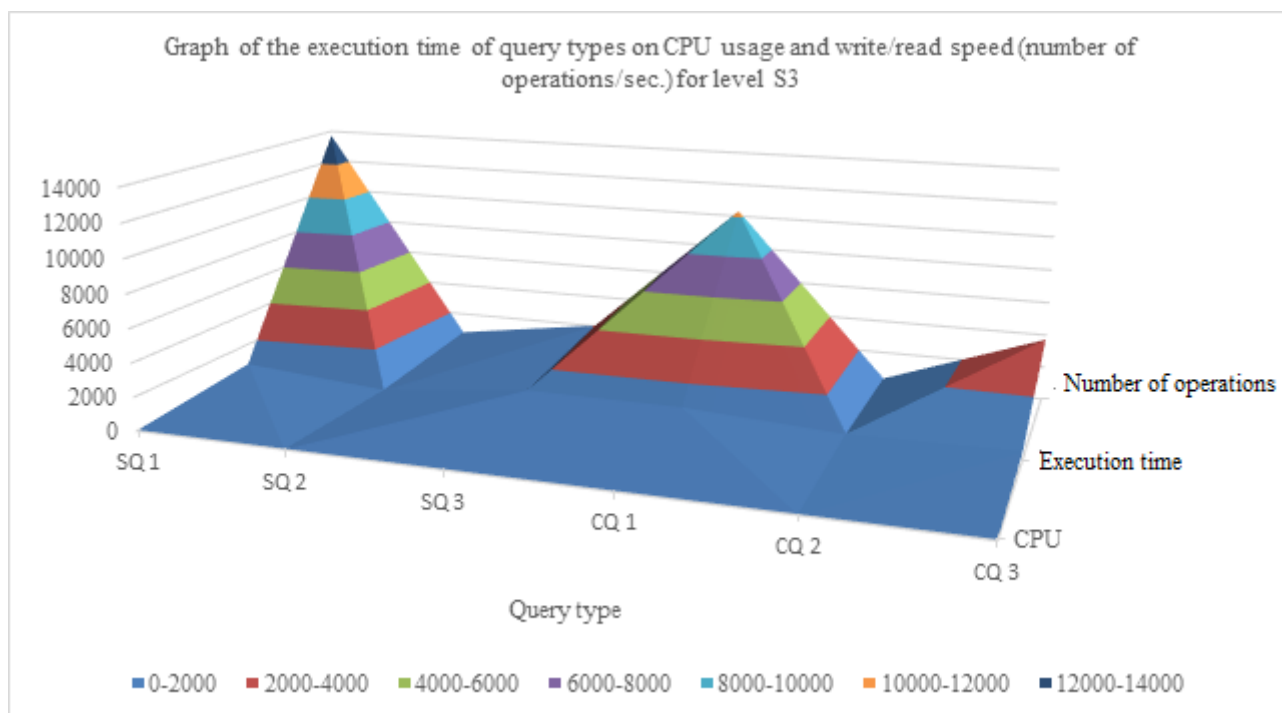


Fig. 3. Graph of query execution time dependence on CPU load and write/read operation speed (number of operations/sec.) for level S3

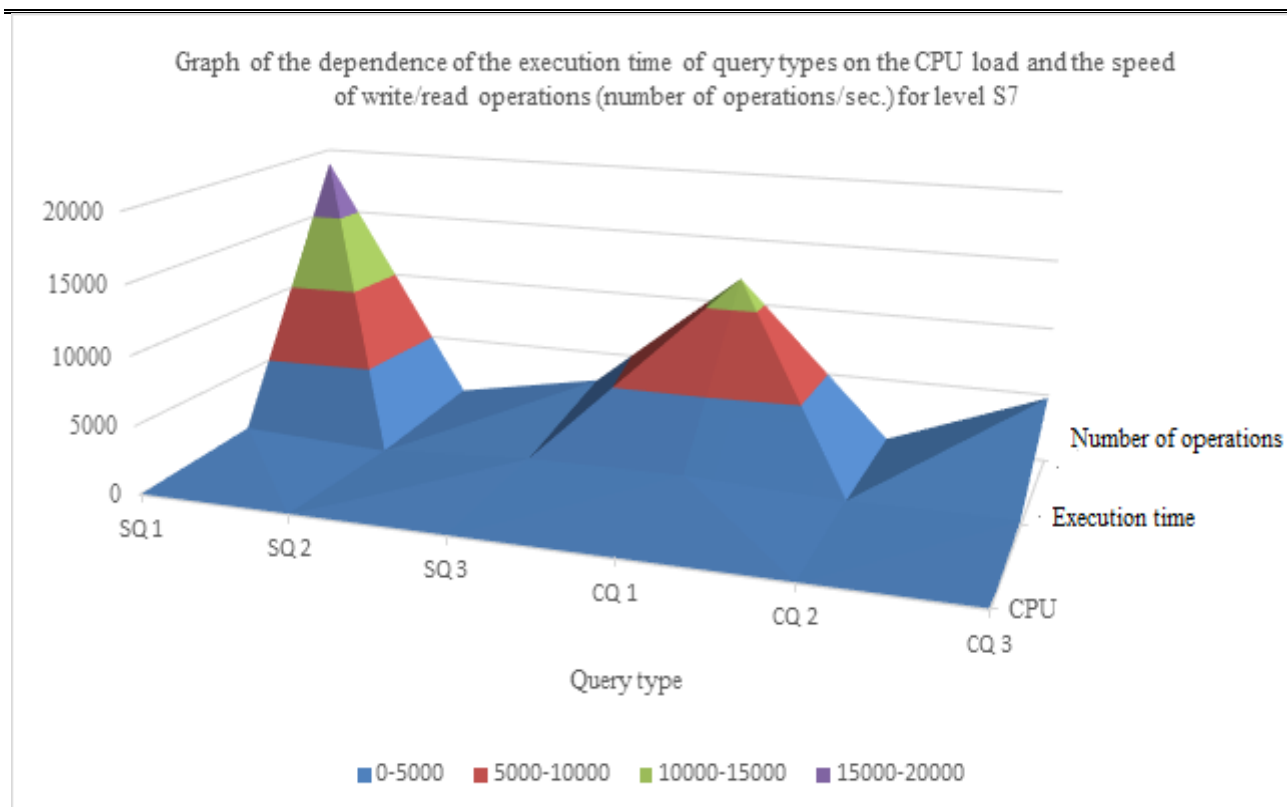


Fig. 4. Graph of dependence of time of execution of requests on CPU loading and speed of write/read operations (number of operations/sec.) for level S7

Conclusions.

The analysis of services and models of database work using cloud platforms showed that in the current global trends of their use the most acceptable from a practical point of view are the models of MS Azure platform – DTU and vCore. To assess the effectiveness of their use in working with relational databases, tools are used to monitor indicators - query execution time and resources involved (CPU and RAM load). It is proved that for a detailed assessment and further analysis of productivity it is necessary to expand the composition of these indicators, for which 2 groups of indicators of database work were proposed for the first time: indicators of execution time and performance of simple and complex queries the key indicator is the level of the DTU model. An additional factor influencing the performance of the selected DTU model is the data generation time – the

results show that the generation of large data (10 GB or more) requires the use of the S7 level. Thus, based on the results of the experiments, we can draw a general conclusion – the optimal choice is the level of the model within S3 and S7. The developed set of database performance indicators on the Azure cloud platform expands the basis of methodological principles of using and evaluating the performance of relational databases on cloud platforms [16 - 19] by analyzing the results of simple and complex queries for the DTU model and allows final selection of a model key factors - the type of queries and the amount of test data.

Practical use of the proposed indicators for the DTU model will increase the efficiency of decision-making on the choice of model level in the implementation of different types of queries and database data generation on cloud platforms.

References

1. TOP 10 cloud trends in 2018, available at: <https://ecm-journal.ru/post/TOP-10-oblachnykh-tendencii-v-2018-godu.aspx>.
2. What is cloud computing?, available at: <https://aws.amazon.com/ru/what-is-cloud-computing/>.
3. Gartner., available at: <https://www.gartner.com/en>.
4. Marinescu, D., C. (2018), "Cloud Service Providers and the Cloud Ecosystem", *Cloud Computing*, No. 2, P. 13 – 49. DOI: <https://doi.org/10.1016/B978-0-12-812810-7.00002-9>
5. Database prices SQL Azure, available at: <https://azure.microsoft.com/ru-ru/pricing/details/sql-database/single/>.
6. Database SQL Azure, available at: <https://azure.microsoft.com/ru-ru/services/sql-database/>.
7. "How to migrate to Amazon RDS for SQL Server using transactional replication", available at: <https://aws.amazon.com/ru/blogs/database/how-to-migrate-to-amazon-rds-for-sql-server-using-transactional-replication/>.
8. Amazon RDS pricing for SQL Server, available at: <https://aws.amazon.com/ru/rds/sqlserver/pricing/>.
9. Delaney, K. (2018), *Microsoft Azure Options for SQL Server Relational Databases*, Microsoft Corporation, 22 p.

10. Marozzo, F. (2019), "Infrastructures for High-Performance Computing: Cloud Infrastructures", *Reference Module in Life Sciences*, No. 1, P. 240 – 246. DOI: <https://doi.org/10.1016/B978-0-12-809633-8.20374-9>
11. Boukhelef, D., Boukhobza, J., Boukhalfa, K., Ouarnoughi, H., Lemarchand, L. (2019), "Optimizing the cost of DBaaS object placement in hybrid storage systems", *Future Generation Computer Systems*, Vol. 93, P. 176 – 187. DOI: <https://doi.org/10.1016/j.future.2018.10.030>
12. Saif, S., Wazir, I. (2018), "Performance Analysis of Big Data and Cloud Computing Techniques: A Survey", *Procedia Computer Science*, No. 132, P. 118 – 127. DOI: <https://doi.org/10.1016/j.procs.2018.05.172>
13. Importing and Exporting SQL Server Databases, available at: https://docs.aws.amazon.com/en_us/AmazonRDS/latest/UserGuide/SQLServer.Procedural.Importing.html.
14. Resource limits for single databases using the DTU-based purchasing model, available at: <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-dtu-resource-limits-single-databases>.
15. Azure SQL Database DTU Calculator, available at: <http://dtucalculator.azurewebsites.net/>.
16. Passos da Costa, P. J., Rosado da Cruz, A. M. (2012), "Migration to Windows Azure – Analysis and Comparison", *Procedia Technology*, Vol. 5, P. 93–102. DOI: <https://doi.org/10.1016/j.protcy.2012.09.011>
17. Han, J., Kamber, M., Pei, J. (2012), "Data Warehousing and Online Analytical Processing". DOI: <https://doi.org/10.1016/B978-0-12-381479-1.00004-6>
18. Dabbechi, H., Nabli, A., and Bouzguenda, L. (2016), "Towards Cloud-Based Data Warehouse as a Service for Big Data Analytics", *Lecture Notes in Computer Science*, 9876, P.180–189. DOI: <https://doi.org/10.1007/978-3-319-45246-317>
18. Zolotariov, D. (2020), "The system of automated billing on the basis of bad infrastructure has been developed" ["Rozpodilena sistema avtomatizovanih obchislen' na bazi hmarnoi infrastrukturi"], *Innovative technologies and scientific solutions for industry*, No. 4 (14), P.47-56. DOI: <https://doi.org/10.30837/ITSSI.2020.14.047>

Received 14.01.2022

Відомості про авторів / Сведения об авторах / About the Authors

Мінухін Сергій Володимирович – доктор технічних наук, професор, професор кафедри інформаційних систем ХНЕУ ім. С.Кузнеця, Харків, Україна; e-mail: minukhin.sv@gmail.com; ORCID: <https://orcid.org/0000-0002-9314-3750>.

Минухин Сергей Владимирович – доктор технических наук, профессор, профессор кафедры информационных систем ХНЭУ им. С.Кузнеця, Харьков, Украина.

Minukhin Serhii – Doctor of Sciences (Engineering), Professor, Professor of the Department of Information Systems, Simon Kuznets Kharkiv National University of Economics Kharkiv, Ukraine.

ДОСЛІДЖЕННЯ ПРОДУКТИВНОСТІ МОДЕЛІ DTU ДЛЯ РЕЛЯЦІЙНИХ БАЗ ДАНИХ НА ПЛАТФОРМІ AZURE

При рішенні завдань роботи з реляційними БД на хмарних платформах виникає проблема вибору певної моделі для забезпечення продуктивності виконання запитів різної складності. **Об'єктом** дослідження є процеси реалізації різних типів запитів до реляційних БД в рамках моделі придбання DTU платформи MS Azure. **Предметом** є методи оцінювання продуктивності роботи з реляційними БД на основі часових показників виконання запитів та показників завантаженості ресурсів хмарної платформи. **Мета** дослідження полягає в розробленні системи показників для моніторингу поточного стану роботи з БД для обґрунтованого прийняття рішень щодо вибору певної цінової категорії моделі DTU хмарного сервісу MS Azure, що дозволить оптимізувати результати роботи з БД платформи. Досягнення поставленої мети передбачає виконання таких **завдань**: проаналізувати сучасні засоби та сервіси роботи з БД, зокрема реляційними БД, на хмарних платформах Azure та AWS, особливості їх застосування та реалізації; розробити програмне забезпечення для генерації тестових реляційних БД різних об'ємів; провести тестування згенерованих БД на локальному ресурсі; з врахуванням характеристик рівнів моделі DTU Azure розробити нову систему показників продуктивності, яка включає 2 групи - часові показники та показники завантаженості існуючих ресурсів платформи; розробити та реалізувати запити різної складності для згенерованої тестової БД для різних рівнів моделі DTU та провести аналіз отриманих результатів. **Методи**. В дослідженні використано такі методи: методи проектування реляційних баз даних; методи створення запитів у SQL-орієнтованих базах даних з довільною кількістю таблиць; методи створення та міграції даних у хмарні платформи; методи моніторингу результатів виконання запитів на основі часових та ресурсних показників; методи генерації тестових даних для реляційних БД; системний підхід для комплексного оцінювання та аналізу продуктивності роботи з реляційними БД. **Результати**. На основі розробленої системи показників, що використовується для поточного аналізу процесів роботи з реляційними БД платформи MS Azure, проведено чисельні експерименти для різних рівнів моделі для простих та складних запитів до БД загальним об'ємом 20 ГБ: навантаження рівнів моделі DTU при виконанні різних запитів, вплив рівнів моделі DTU Azure SQL database на показники виконання простих та складних запитів, залежність часу виконання різних запитів від завантаженості CPU (ЦП) та швидкості операцій запису/читання для різних рівнів моделі. **Висновки**. Отримані результати експериментів дозволяють зробити висновок щодо використання рівнів моделі DTU - S3 та S7 - для генерації тестових даних різного об'єму (до 20 ГБ) та виконання запитів до БД. Практичне використання запропонованих показників для оцінювання результатів застосування моделі DTU дозволить підвищити ефективність прийняття рішень щодо вибору рівня моделі при реалізації різних запитів та генерації тестових даних БД на хмарній платформі Azure. Розроблений набір показників роботи з реляційними БД

на хмарній платформі Azure розширює базис методологічних засад оцінювання продуктивності роботи з реляційними БД на хмарних платформах шляхом аналізу результатів виконання простих та складних запитів до БД на заданих ресурсах.

Ключові слова: хмарна платформа; реляційна база даних; модель придбання DTU; показники часу та завантаженості; генератор даних; тестові дані; складність запиту.

ИССЛЕДОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ МОДЕЛИ DTU ДЛЯ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ НА ПЛАТФОРМЕ AZURE

При решении задач работы с реляционными БД на облачных платформах возникает проблема выбора определенной модели для обеспечения производительности выполнения запросов разной сложности. **Объектом** исследования являются процессы реализации различных типов запросов к реляционным базам данных в рамках модели приобретения DTU платформы MS Azure. **Предметом** является методы оценки производительности работы с реляционными базами данных на основе временных показателей выполнения запросов и показателей загрузки ресурсов облачной платформы. **Цель** исследования заключается в разработке системы показателей для мониторинга текущего состояния работы с БД для обоснованного принятия решений выбора определенной ценовой категории модели DTU облачного сервиса MS Azure, что позволит оптимизировать результаты работы с БД. **Достижение** поставленных целей предполагает выполнение следующих **задач**: проанализировать современные средства и сервисы работы с БД, в частности реляционными БД, на облачных платформах Azure и AWS, особенности их применения и реализации; разработать программное обеспечение для генерации тестовых реляционных БД разных объемов; провести тестирование сгенерированных БД на локальном ресурсе; с учетом характеристик уровней модели DTU Azure разработать новую систему показателей производительности, которая включает 2 группы – временные показатели и показатели загрузки существующих ресурсов платформы; разработать и реализовать запросы разной сложности для сгенерированной тестовой БД для разных уровней модели DTU и провести анализ полученных результатов. **Методы.** В исследовании использованы следующие методы: методы проектирования реляционных баз данных; методы создания запросов в SQL-ориентированных базах данных с произвольным количеством таблиц; методы создания и миграции данных в облачные платформы; методы мониторинга результатов выполнения запросов на основе временных и ресурсных показателей; методы генерации тестовых данных для реляционных БД; системный подход для комплексной оценки и анализа производительности работы с реляционными БД. **Результаты.** На основе разработанной системы показателей, используемой для текущего анализа процессов работы с реляционными БД платформы MS Azure, проведены многочисленные эксперименты для разных уровней модели для простых и сложных запросов к БД общим объемом 20 ГБ: нагрузка уровней модели DTU при выполнении различных запросов, влияние уровней модели DTU Azure SQL database на показатели выполнения простых и сложных запросов, зависимость времени выполнения различных запросов от загрузки CPU (ЦП) и скорости операций записи/чтения для разных уровней модели. **Выводы.** Полученные результаты экспериментов позволяют сделать вывод об использовании уровней модели DTU - S3 и S7 - для генерации тестовых данных разного объема (до 20 ГБ) и выполнения запросов к БД. Практическое использование предложенных показателей для оценки результатов применения модели DTU позволит повысить эффективность принятия решений по выбору уровня модели при реализации различных запросов и генерации тестовых данных на облачной платформе Azure. Разработанный набор показателей работы с реляционными БД на облачной платформе Azure расширяет базис методологических основ для оценки производительности работы с реляционными БД на облачных платформах путем анализа результатов выполнения простых и сложных запросов к БД на задействованных ресурсах.

Ключевые слова: облачная платформа; реляционная база данных; модель приобретения DTU; показатели времени и загрузки; генератор данных; тестовые данные; сложность запроса.

Бібліографічні описи / Bibliographic descriptions

Мінухін С. В. Дослідження продуктивності моделі DTU для реляційних баз даних на платформі Azure. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 27–39. DOI: <https://doi.org/10.30837/ITSSI.2022.19.027>

Minukhin, S. (2022), "Performance study of the DTU model for relational databases on the Azure platform", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 27–39. DOI: <https://doi.org/10.30837/ITSSI.2022.19.027>

A. POPOV, M. MOMOT, A. YELIZIEVA

CHOOSING THE TEST AUTOMATION SYSTEM ACCORDING TO CUSTOMER REQUIREMENTS

The **subject** of the research are methods and technologies for automating the process of software product testing. The aim of the work is to optimize the time and costs for performing automated testing of software products. The following **tasks** were solved: analysis of existing software testing automation systems; formation of system of selection criteria for testing automation systems; development of formalized model of selection process; development of automation system selection algorithm considering customer's requirements; development of UML diagrams for presentation of functional capabilities of developed application; development of application for informational support of selection process. To solve these **tasks**, we used methods of system analysis, theory of sets and technologies of cross-platform applications development. The following **results** were obtained: The most popular systems of test automation have been analyzed, their scope and main capabilities have been singled out. Selection criteria are singled out, divided into qualitative and quantitative. Formalized model for choosing test automation systems taking into account their characteristics and customer requirements is proposed. Developed UML diagram shows the functionality of the developed subsystem. The proposed algorithm for determining the re-recommended system of test automation allows us to take into account the vectors of criteria for testing systems. On the basis of the formalized model and algorithm we developed a subsystem that allows us to determine the optimal variant of test automation system on the basis of the introduced selection criteria. **Conclusions:** informational support for choosing a test automation system for software products based on the developed algorithm takes into account the customer's requirements and the characteristics of the existing systems, which allows us to select the most preferable option out of the possible systems. The main result of the developed subsystem is a recommendation for a user to use an automated testing system, taking into account customer requirements.

Keywords: testing of the software; automation testing systems; criteria of choice of the systems; the model of multicriteria choice; the algorithm of choice of the system; a diagram of precedents.

Introduction

When planning work on the development of a software system, much attention is paid to one of the stages of the life cycle - testing. Testing is a very important stage because it is at this stage the developed system takes the form of a full-fledged product. At present, systematized testing is actively used in the industrial development of software products. Automation of this process becomes more and more necessary with increasing scale of projects [1]. Automatization systems significantly simplify the process and save both time and money [2, 3].

At present, there is a large number of test automation systems on the market, both traditionally popular and newly emerging. A software company faces a nontrivial task of selecting test automation systems, taking into account existing restrictions/requirements, such as supported testing processes, technologies, standards, integration with development systems, application type, platform type, test execution time, budgets

When implementing the testing phase, it is necessary to meet the resource conditions of the software product development project [4, 5]. It is necessary to take into

account the time and cost constraints, which may be the requirements of the customer.

Thus the purpose of this article is to optimize the time and cost of performing automated testing. The article presents analysis and comparison of existing testing automation systems, development of a mathematical model for choosing the automatic testing subsystem and an algorithm of the subsystem performance for choosing an appropriate system for automating the testing process.

Analysis of recent research and publications

The process of automatic testing involves checking the software product for various inaccuracies and errors that were made at the development stage [6]. Analysis of publications on the subject showed that the main attention is paid to testing methods (functional, system, modulus, performance testing, etc.) and capabilities of software tools [6, 7]. Nowadays there are many different automation systems, which are used at testing stage of software products [8, 9]. Table 1 presents a comparative analysis of existing software testing automation systems, reveals their main features and application features [10-17].

Table1. Analysis of existing test automation systems

Name	Field of application	Features
1	2	3
Selenium	It is used to automate the testing of web applications, as well as any routine actions performed through the browser. Selenium supports desktop and mobile browsers.	<ul style="list-style-type: none"> - allows you to develop automation scripts in almost any programming language; - the ability to organize distributed stands that consist of hundreds of machines with different operating systems and browsers; - the ability to run scenarios in the clouds.

The end **Table1.**

1	2	3
TestComplete	Allows you to create tests for Windows applications, web servers and web pages. Used to automate various types of software testing for .NET, Java, Visual C++, Visual Basic, Delphi, C++Builder, web pages and other applications in order to reduce the time required for manual testing.	<ul style="list-style-type: none"> - the presence of a special mechanism that facilitates the creation of scripts for testing applications; - the necessary actions of the tester are automatically recorded in a file that is available for viewing and editing and repeats all previously performed operations when launched; - recorded tests can be modified later.
Quick Test Professional (QTP)	Faster than its counterparts recognizes graphic controls ("controls")	<ul style="list-style-type: none"> - offers a wide range of functionality for checking and interacting with the application under test; - uses the VBScript programming language to automate tests; - offers several types of checkpoints that allow users to control different aspects of application testing.
MS Test	Requires installation of Visual Studio, as it is part of the Team Foundation Server product. Plans and test results are saved on the Team Foundation Server. Test scenarios are managed by Microsoft Test Manager (MTM). MTM includes a test plan, test case, and configurations. MTM works only at the application level, so it needs to be installed on the server (if the server is remote, the work is done via VPN). Provides the ability to combine tasks that are assigned to the performer, with the reports (defects) and reports on the time spent on the work.	<ul style="list-style-type: none"> - research testing; - planning and execution of manual tests; - cross-platform test configurations (different versions of the same test for different platforms/releases); - test passing diagnostics (logs, videos, etc.); - import-export of tests; - cross-project import/export of tests; - recording and playback of manual tests (recorder); - test automation.
Telerik WebUI Test Studio	Used to automate functional testing of web applications (ASP.NET, Silverlight). It is an extension to Microsoft Visual Studio	<ul style="list-style-type: none"> - Silverlight support. In the initial versions of the program there were enough bugs associated with the support of silverlight elements, now the problem has been eliminated; - excellent handling of AJAX requests of any complexity; - no problems when writing tests with RadControls components (page controls); - has a handy test editor; - integration with other testing tools. All features of Visual Studio are available; - basic functionality of any sensible modern automation tool (Test Recorder, Elements Explorer, DOM Explorer, code generation, support for multiple browsers, etc.) is implemented at a good level.

The effectiveness of the testing phase depends on the correct choice of an automation tool. The choice of an application test automation system depends on the following factors:

- the type of application to be tested (platform orientation);
- the testing objectives (compliance with standards, applicability of the system);
- the requirements of the client (deadline and cost).

Based on this, there is a need to develop a formalized model for selecting a means of performing automated testing.

Materials and methods

To formalize the process of selecting a system for automating the testing process, the following model is proposed. Let's represent in the form of a vector of sets of criteria affecting the choice of a test automation system:

$$K = \{k_i\}, i = \overline{1,6}. \quad (1)$$

Two types of criteria are used: quantitative and qualitative [18]. Quantitative criteria include:

- K_1 – timing (necessary time to carry out the work);
- K_2 – cost of the work (the cost of their implementation);

- K_3 – number of employed specialists.

Qualitative criteria include:

- K_4 – degree of applicability of the system;
- K_5 – degree of compliance with standards;
- K_6 – platform orientation.

The degree of applicability of systems allows you to determine how effectively a test automation system can be adapted to the real work on a project when going through the testing phase, and which of its benefits can be most useful in doing so [19, 20]. The degree of compliance with standards shows to what extent the various characteristics

of activities in the testing phase comply with international standards in this area:

- ISO 90003:2014 (Guidelines for the application of ISO 9001:2008 to computer software);
- ISO 25051:2014 (Requirements for quality of Ready to Use Software Product (RUSP) and instructions for testing);
- IEEE 829 (Software test documentation);
- IEEE 1008 (Standard for Software Unit Testing);
- ISO 8402 (Quality management and quality assurance);
- ISO 9126 (Software engineering – Product quality).

Each of the standards has its own specific specification and reflects a particular aspect of the testing phase activity. The degree of conformance to the standards shows how well the testing phase has been prepared, depending on each specific standard. For example, IEEE Standard 829 describes the procedure for compiling a test plan that will be used for the work, IEEE Standard 1008 describes the procedure for organizing unit testing, etc.

It should be noted that the choice of test automation system is carried out on at least two criteria K_i , which are most important to the customer. Let's represent the existing test automation systems as a set M [21, 22]:

$$M = \{m_j\}, j = \overline{1, N}. \quad (2)$$

In this case, the values of the vector of criteria for each test automation system are known:

$$\forall m_j \in M : K = \langle k_i(m_j) \rangle. \quad (3)$$

It should be noted that these criteria are heterogeneous. To bring them into a dimensionless value, it is necessary to normalize them by the formula:

$$k_i^H(m_j) = \frac{k_i(m_j) - k_i^{Wx}(m_j)}{k_i^{Hx}(m_j) - k_i^{Wx}(m_j)}, \quad (4)$$

where $k_i^{Hx}(m_j)$ – the best value of the criterion;

$k_i^{Wx}(m_j)$ – the worst value of the criterion;

$k_i(m_j)$ – the current criterion values.

The value of quantitative criteria is represented by some non-negative number. Qualitative criteria have point estimates, i.e. they are also represented by non-negative numbers. These evaluations can be obtained by expert way. Experts also determine the degree of importance of criteria, i.e. a tuple of dimensionless weight coefficients can be determined:

$$A = \langle a_i \rangle, i = \overline{1, n}.$$

The condition must be met that:

$$0 \leq a_i \leq 1, \forall i = \overline{1, n}, \sum_{i=1}^n a_i = 1. \quad (5)$$

Thus, we can introduce a utility function:

$$P(x) = F[A, k_i^H(m_j)], \quad (6)$$

where $k_i^H(m_j)$ – normalized criteria in the interval $[0, 1]$ (4);

A – tuple of dimensionless weight coefficients.

Customer requirements for selecting a test automation system are described by a vector of desired criteria values:

$$R = \{r_i\}, i = \overline{1, 6}. \quad (7)$$

In this case, if the customer is not interested in some of the criteria, the corresponding criteria in the vector can take the value -1 (as a negative signal value).

It is also necessary to fulfill the condition of resource feasibility of the testing stage:

$$res_k \leq Res_l, \quad (8)$$

where Res_l – l -th resource (time, financial, etc.) of the software project;

res_k – k -th type of resource at the testing stage.

The problem statement is formulated as follows: it is necessary to form the resulting set $M_{rec} \subset M$, which consists of alternatives, corresponding in their values of criteria (3) to the requirements of the customer:

$$M_{rec} = \{(k_i(m_j)) | (k_i(m_j) \leq r_i) \cup (k_i(m_j) > -1)\}. \quad (9)$$

Let's represent the algorithm for choosing the automation system as follows:

Step 1. Initialization: resulting set $\mathbf{M}_{rec} = \emptyset$.

Step 2. For all $m_i \in \mathbf{M}$, $i = 1, 2, \dots, N_M$.

2.1. flag = true.

2.2. For all r_j , $j = 1, 2, \dots, N$.

2.2.1. If $r_j == -1$, then switch to 2.2.

2.2.2. If the matching condition is not met $r_j \div k(m_j^i)$,

then flag = false

2.2.3. If ($j == N$) and (flag == true), then

$\mathbf{M}_{rec} = \mathbf{M}_{rec} \cup m_i$.

The designation means the operation of checking the correspondence ($r_j \div z_j^i$) of the value of the i -th criterion of the j -th test automation system to the value of the same i -th criterion determined by the customer. The final choice of one option of test automation system from the \mathbf{M}_{rec} set (if it contains more than one element) is recommended to be made based on the customer's preferences or using the criteria convolution according to the formula (6).

Study results

Based on a formal representation of the system selection process for test automation and a generalized selection algorithm, a prototype subsystem was developed to recommend a test automation system to the customer based on various types of actors. Fig. 1 shows the USE CASE diagram, which shows the relationships between

the actors and precedents of the subsystem under development.

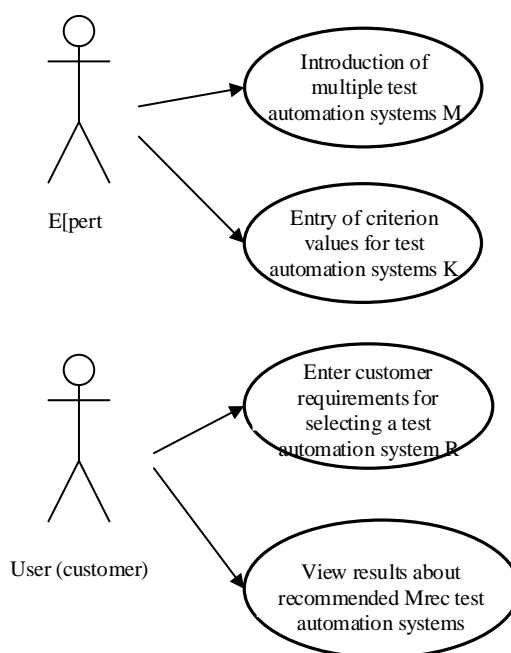


Fig. 1. The USE CASE diagram of the subsystem under development

The input data for the subsystem are the following parameters:

- project name;
- version;
- project type;
- platform used;
- start date of the stage;
- end date;
- estimated duration (days);
- anticipated cost;
- number of specialists employed;
- assessment of compliance with standards (1 to 10), such as: ISO 90003:2014, ISO 90003:2014, ISO 25051:2014, IEEE 829, IEEE 1008, ISO 8402, ISO 9126.

A score of 1 to the standard defines the worst value, and 10 defines the best value, i.e. full compliance of the testing system with the specified software standards. It is

assumed that the subsystem selects the most preferable system for test automation among the systems presented in table 1.

Table 2 shows data on the correspondence of the set of criterion vector values for test automation systems. A combination of parameter values determines the possibility of recommending a particular test automation system.

Based on the values of the previously mentioned criteria k_i , as well as the set M of the considered test automation systems, an algorithm for determining the recommended test automation system was obtained (fig. 2). The incoming arrow index in the system selection block in the algorithm in fig. 2 was used to refine the set of criterion vector values for the systems - written in brackets in the cells of table 2.

Table 2. Values of the criteria vectors for automation systems testing

Values	Selenium			TestComplete				QTP	MS Test		Telerik Test Studio
Application type											
Windows-Application	-1	-1	-1	-1	-1	1(10)	1(11)	-1	-1	1(7)	-1
Windows, .Net	-1	-1	-1	-1	-1	1(10)	-1	-1	-1	1(7)	-1
Windows, Java	-1	-1	-1	-1	-1	-1	1(11)	-1	-1	-1	-1
Web application	1(3)	1(4)	1(5)	1(8)	1(9)	-1	-1	1(2)	1(6)	-1	1(1)
Web, .Net	1(3)	-1	-1	-1	1(9)	-1	-1	-1	1(6)	-1	1(1)
Web, Java	-1	1(4)	-1	1(8)	-1	-1	-1	-1	-1	-1	-1
Web, other	-1	-1	1(5)	-1	-1	-1	-1	1(2)	-1	-1	-1
Budget											
<=10.000	1(3)	-1	-1	-1	-1	1(10)	-1	-1	-1	-1	-1
>10.000	-1	-1	-1	-1	-1	-1	-1	-1	-1	1(7)	-1
<=20.000	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1(1)
>20.000	-1	-1	-1	-1	1(9)	-1	-1	-1	1(6)	-1	-1
Terms											
> 14 days	-1	-1	-1	1(8)	1(9)	-1	-1	1(2)	1(6)	-1	1(1)
<= 14 days	1(3)	1(4)	1(5)	-1	-1	-1	-1	-1	-1	-1	-1

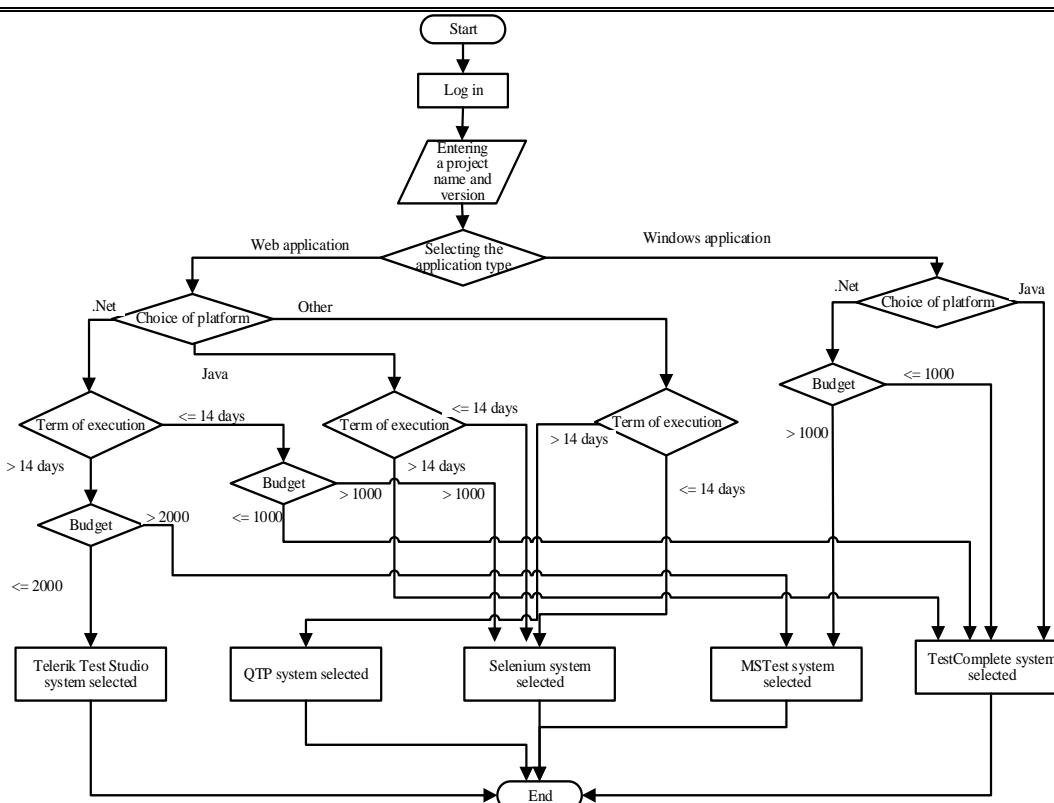


Fig. 2. Algorithm of determination of the recommended test automation system

In the presented scheme, when choosing a web application, the timing and budget are taken into account, because the available systems allow more flexibility in organizing the process of test automation in web development. When choosing a Windows application, the timing and budget (except for the .Net platform) are not taken into account, because the systems supporting these platforms do not allow such flexibility in organizing the automation process.

Generally speaking, the algorithm can be represented by the following steps:

- logging into the system;
- entering the name and version of the project;
- selection of the application type (web application selected);
- entering the start and end date of the stage;
- input of the estimated duration;
- entry of estimated costs.

Consider an example of calculating the choice of the system according to the specified initial data:

Project name: Randomator.

Version: 2.1.2.

Type of the project: Web application.

Platform used: .NET.

Start date of the stage: 25.08.2019.

End date: 12.10.2019.

Estimated duration: 48 days.

Estimated cost: 24000.

Conclusions

A model of multi-criteria choice of test automation system on the basis of the vector of quantitative and qualitative criteria is proposed, which allows taking into account the requirements of the customer and the resources of the project to create a software product with the help of the utility function. The paper presents an algorithm for determining the recommended system of test automation, which is implemented in a prototype software subsystem in which the user receives a specific recommendation for the use of the most effective system of test automation based on the entered data.

Scientific novelty of the article consists in the formalized representation of the process of choosing the system for automating the testing of software products, which allows scientifically justified choice of the system for evaluating the compliance of the finished program with the specified requirements.

Practical significance consists in the possibility of using the developed subsystem for information support of testers of software products. This allows you to choose the best option for automating the testing process, taking into account the requirements of the customer.

References

1. Dustin, E., Garrett, T., Gauf, B. (2009), *Implementing automated software testing: how to save time and lower costs while raising quality*, Addison-Wesley, Boston, 368 p.
2. Berkun, S. (2014), *Art of IT-project management [Iskusstvo upravleniya IT-proektami]*, SPB, St. Petersburg, 700 p.

3. Novichkov, A., Pankratov, V. (2005), "Automated testing: evaluation of investments return and associated risks" ["Avtomatizirovannoe testirovanie: ocenka vozvrata investitsij i soputstvujushhie riski"], *ComputerPress*, No. 11. P. 47 – 49.
4. Golyan, V. V., Kravchenko, O. K. (2019), "Comparison of software life cycle models in order to identify the most effective" ["Porivnyannyya modelej zhyttyevykh cykliv programnogo zabezpechennya z metoyu vyyavleniya najefektyvnishogo"], *Infocommunication systems*, No. 2 (157), P. 63 – 70. DOI: <http://doi.org/10.30748/soi.2019.157.08>
5. Kolesnikova, K. V., Gogunskii, V. D., Negri, A. A., Olekh, H. S. (2016), "Conceptual model of project management" ["Konceptualna model upravlinnya proektamy"], *Electro technical and computers systems*, No. 23 (99), P. 175 – 179. DOI: <https://doi.org/10.15276/eltecs.23.99.2016.27>
6. Yegorova, O., Bychok, V. (2019), "Software testing tools" ["Programni zasoby dlya testuvannya programnogo zabezpechennya"], *Young scientist*, No. 11 (75), P. 680 – 684. DOI: <https://doi.org/10.32839/2304-5809/2019-11-75-144>
7. Ivan Romero-Pena, Gerardo Padilla-Zarate, Karen Cortes-Verdin (2021), "Identification of Test Cases Duplication: Systematic Literature Review. Software Engineering Research and Innovation", *9th International Conference*, P. 104 – 111.
8. Troyan, A. M., Modenov, Yu. B. (2017), "The advisability of automated testing to ensure the quality of the software products" ["Docilnist avtomatyzovanogo testuvannya dlya zabezpechennya yakosti programnykh produktiv"], *Problems of automatization and control*, Vol. 1, No. 57 – 58, P. 86 – 89. DOI: <https://doi.org/10.18372/2073-4751.1.12798>
9. "Top 10 of software test automation tools", available at: <https://habr.com/ru/post/481294/> (last accessed: 09.10.2021).
10. "Top 10 tools of test automation 2018", available at: <https://habr.com/ru/post/342234/> (last accessed: 09.10.2021).
11. Slinkin, A., Demianov, Ye. (2010), *Selenium. Introduction to the system [Selenium. Vvedenie v sistemu]*, Simvol-Pljus, Moscow, 656 p.
12. Gundecha, U. (2012), *Selenium Testing Tools Cookbook*, Packt Publishing, Birmingham, 326 p.
13. Alpaev, G. (2013), *TestComplete Cookbook*, Packt Publishing, Birmingham, 282 p.
14. Alpaev, G., *TestComplete Book*, available at: <https://alpaev.com/testcomplete/> (last accessed: 27.10.2021).
15. Sankar, V. R. (2012), *QuickTest Professional*, Tata McGraw-Hill, New Delhi, 600 p.
16. "Using MSTest platform in module tests" ["Ispolzovanie platformy MSTest v modul'nykh testah"], available at: <https://docs.microsoft.com/ru-ru/visualstudio/test/using-microsoft-visualstudio-testtools-unittesting-members-in-unit-tests> (last accessed: 27.10.2021).
17. "Telerik Test Studio", available at: <https://docs.telerik.com/teststudio> (last accessed: 29.10.2021).
18. Mukhamediev, R. I., Mustakayev, R., Yakunin, K., Kiseleva, S., Gopejenko, V., "Multi-Criteria Spatial Decision Making Supportsystem for Renewable Energy Development in Kazakhstan," available at: <https://s3-us-west-2.amazonaws.com/ieeeshutpages/xplore/xplore-ie-notice.html> (last accessed: 27.12.2021).
19. Hochbaum, D. (1995), "Approximating Covering and Packing Problems", *Journal of the Association for Computing Machinery*, Vol. 32, No. 1, P. 94 – 143.
20. Mukhamediev, R. I., Mustakayev, R., Yakunin, K. O., Kuchin, Y. I., Kiseleva, S. V., Gopejenko, V.I. (2020), "Decision Support System for Optimization of Res Generators Placement Based on Geospatial Data", *News of The National Academy of Sciences of The Republic of Kazakhstan Series of Geology and Technical Sciences*, Vol. 1, No. 439, P. 81 – 89. DOI: <https://doi.org/10.32014/2020.2518-170X.10>
21. Beskorovainyi, V. (2020), "Combined method of ranking options in project decision support systems", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (14), P. 13 – 20. DOI: <https://doi.org/10.30837/ITSSI.2020.14.013>
22. Malyyeva, O. V., Yelizieva A., V., Kosenko, N. V., Nevlyudova, V. V. (2018), "Information technology of decision making support on production enterprise purchasing management" ["Informacijna tehnologiya pidtrymky pryjnyattya rishen z upravlinnya zakupivlyamy vyrobnychogo pidpryyemstva"], *Innovative Technologies and Scientific Solutions for Industries*, No. 3 (5), P. 57 – 66. DOI: <https://doi.org/10.30837/2522-9818.2018.5.057>

Received 04.02.2022

Відомості про авторів / Сведения об авторах / About the Authors

Попов Андрій Вячеславович – кандидат технічних наук, доцент, Національний аерокосмічний університет ім. М.Є. Жуковського "ХАІ", доцент кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: apv@xai.edu.ua; ORCID: <https://orcid.org/0000-0001-8984-731X>.

Попов Андрей Вячеславович – кандидат технических наук, доцент, Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", доцент кафедры компьютерных наук и информационных технологий, Харьков, Украина.

Popov Andrei – PhD (Technical Sciences), Docent, National Aerospace University "Kharkiv Aviation Institute"; Docent of the Department of Computer Sciences and Information Technologies, Kharkiv, Ukraine.

Момот Мирослав Олександрович – кандидат технічних наук, доцент, Національний аерокосмічний університет ім. М.Є. Жуковського "ХАІ", доцент кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: m.momot@khai.edu; ORCID: <https://orcid.org/0000-0003-2580-5908>.

Момот Мирослав Александрович – кандидат технических наук, доцент, Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", доцент кафедры компьютерных наук и информационных технологий, Харьков, Украина.

Momot Myroslav – PhD (Technical Sciences), Docent, National Aerospace University "Kharkiv Aviation Institute"; Docent of the Department of Computer Sciences and Information Technologies, Kharkiv, Ukraine.

Єлізева Аліна Володимирівна – кандидат технічних наук, Національний аерокосмічний університет ім. М.Є. Жуковського "ХАІ", доцент кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: a.elizeva@khai.edu; ORCID: <https://orcid.org/0000-0002-8228-9383>.

Елизева Алина Владимировна – кандидат технических наук, Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", доцент кафедры компьютерных наук и информационных технологий, Харьков, Украина.

Yelizieva Alina – PhD (Technical Sciences), National Aerospace University "Kharkiv Aviation Institute"; Docent of the Department of Computer Sciences and Information Technologies, Kharkiv, Ukraine.

ВЫБОР СИСТЕМЫ АВТОМАТИЗАЦИИ ТЕСТИРОВАНИЯ С УЧЕТОМ ТРЕБОВАНИЙ ЗАКАЗЧИКА

Предметом исследования являются методы и технологии автоматизации процесса тестирования программных продуктов. **Целью** работы является оптимизация времени и затрат на выполнение автоматического тестирования программного продукта. В работе решены следующие **задачи**: проведение анализа существующих систем автоматизации тестирования программных продуктов; формирование системы критериев выбора систем автоматизации тестирования; разработка формализованной модели процесса выбора; разработка алгоритма выбора системы автоматизации с учетом требований заказчика; разработка UML диаграмм для представления функциональных возможностей разработанного приложения; разработка приложения для информационной поддержки процесса выбора. Для решения указанных задач были использованы **методы** системного анализа, теории множеств, технологии разработки кросс-платформенных приложений. Получены следующие **результаты**. Проанализированы наиболее популярные системы автоматизации тестирования, выделены их область применения и основные возможности. Выделены критерии выбора, которые разделяются на количественные и качественные. Предложена формализованная модель выбора систем автоматизации тестирования с учетом их характеристик и требований заказчика. Разработанная UML диаграмма отражает функциональные возможности разрабатываемой подсистемы. Предложенный алгоритм определения рекомендуемой системы автоматизации тестирования позволяет учитывать векторы критериев для систем тестирования. На основе формализованной модели и алгоритма была разработана подсистема, которая на основе введенных критериев выбора позволяет определить оптимальный вариант системы автоматизации тестирования. **Выводы**: информационная поддержка выбора системы автоматизации тестирования программных продуктов на основе разработанного алгоритма учитывает требования заказчика и характеристики существующих систем, что позволяет выбрать наиболее предпочтительный вариант из возможных систем. Основным результатом работы разработанной подсистемы является получение пользователем рекомендации по использованию системы автоматического тестирования с учетом требований заказчика.

Ключевые слова: тестирование программ; системы автоматизации тестирования; критерии выбора системы; модель многокритериального выбора; алгоритм выбора системы; диаграмма прецедентов.

ВИБІР СИСТЕМИ АВТОМАТИЗАЦІЇ ТЕСТУВАННЯ З УРАХУВАННЯМ ВИМОГ ЗАМОВНИКА

Предметом дослідження є методи та технології автоматизації процесу тестування програмних продуктів. **Метою** роботи є оптимізація часу й витрат на проведення автоматичного тестування програмного продукту. В роботі вирішені наступні **задачі**: проведення аналізу існуючих систем автоматизації тестування програмних продуктів; формування системи критеріїв вибору систем автоматизації тестування; розробка формалізованої моделі процесу вибору; розробка алгоритму вибору системи автоматизації з урахуванням вимог замовника; розробка UML діаграм для подання функціональних можливостей розробленого додатку; розробка додатку для інформаційної підтримки процесу вибору. Для розв'язання вказаних задач були використані **методи** системного аналізу, теорії множин, технології розробки крос-платформних додатків. Отримані наступні **результати**. Проаналізовані найбільш популярні системи автоматизації тестування, виділені їх область застосування й основні можливості. Виділені критерії вибору, які розділяються на кількісні та якісні. Запропонована формалізована модель вибору систем автоматизації тестування з урахуванням їх характеристик і вимог замовника. Розроблена UML діаграма відображає функціональні можливості розробленої підсистеми. Запропонований алгоритм визначення рекомендованої системи автоматизації тестування дозволяє враховувати вектори критеріїв для систем тестування. На основі формалізованої моделі та алгоритму була розроблена підсистема, яка на основі введених критеріїв вибору дозволяє визначити оптимальний варіант системи автоматизації тестування. **Висновки**: інформаційна підтримка вибору системи автоматизації тестування програмних продуктів на основі розробленого алгоритму враховує вимоги замовника й характеристики існуючих систем, що дозволяє обирати найбільш переважний варіант серед можливих систем. Основним результатом роботи розробленої підсистеми є отримання користувачем рекомендацій з використання системи автоматичного тестування з урахуванням вимог замовника.

Ключові слова: тестування програм; системи автоматизації тестування; критерії вибору системи; модель багатокритеріального вибору; алгоритм вибору системи; діаграма прецедентів.

Бібліографічні описи / Bibliographic descriptions

Попов А. В., Момот М. О., Єлізева А. В. Вибір системи автоматизації тестування з урахуванням вимог замовника. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 40–46. DOI: <https://doi.org/10.30837/ITSSI.2022.19.040>

Popov, A., Momot, M., Yelizieva, A. (2022), "Choosing the test automation system according to customer requirements", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 40–46. DOI: <https://doi.org/10.30837/ITSSI.2022.19.040>

L. SMIDOVYCH, Y. DAVYDOVSKYI

PROCESSES OF THE TELECOM OPERATOR'S INFORMATION ARCHITECTURE TRANSFORMATION

The **subject** of the article is the architecture of information systems of telecom operators (BSS/OSS), requirements for it and the processes of its transformation. The **goal** is to transform the existing information environment of the operator into an open digital ecosystem, in accordance with business and technological requirements for the functionality and architecture of information systems of communication operators, which will introduce new business models, services and products, communication channels, and increase competitiveness. The following **tasks** were solved at the study: analysis of the current state of telecommunications and information systems used by telecom operators; analysis of requirements for business support and operational information systems, collection and analysis of requirements for the architecture of modern information systems in the field of telecommunications, analysis of the transformation process from the existing BSS/OSS architecture of the communication operator into the target architecture of the digital communications provider, elaboration of a recommendations for an intermediate architecture. To solve these problems, **methods** of system and business analysis were used, in particular, methods of analysis of software requirements, technologies for developing software application architecture. The following **results** were obtained. The current state of the telecommunications services industry, the features of the existing information systems of telecommunications operators (BSS/OSS) and the requirements for the architecture of such systems was analyzed. Strategies of transition to modern architecture was considered and analyzed. The target architecture of the digital service provider's information ecosystem and the process of transforming the existing BSS/OSS architecture into the target architecture were considered. Intermediate migration architecture was proposed. **Conclusions:** the transformation to modern information architecture will ensure compliance of BSS/OSS systems with market requirements and new technologies, adapt to their changes, allow operators to increase competitiveness, reduce operating and capital costs, create technological conditions for open digital ecosystem and transform the operator into a digital services provider. It was recommended to use an intermediate architecture based on the open standardized TMF API for such transformation.

Keywords: BSS; OSS; information architecture; telecommunications; telecom operator; requirements analysis; digital service provider; open information ecosystem; omnichannel; architecture transformation processes.

Introduction

Currently, the telecommunications sector worldwide is at the stage of changing technologies and business models, which has been called digital transformation. The driver is both a change in technology - the transition to the digital architecture of the IMS network and 4G/5G technologies, the development of technologies and services IoT/IoE, and the emergence of new models of doing business (including OTT, B2B2x/B2G2x) and changing customer expectations.

Under these conditions, the role of software is growing, which not only ensures the automation of operating and business processes, but also, in fact, the implementation of network functions and communication services.

One of the main components of the information infrastructure of the communication service provider (CSP) is OSS/BSS (Operation Support System/Business Support System). These systems manage interaction with subscribers and partners on the one hand, and network infrastructure and resources - on the other, and also to automate corresponding business processes.

Despite standardization efforts, the information infrastructure of most CSPs, especially those with a long history of operation, is a set of legacy information systems and applications, and has a number of shortcomings, which will be discussed below.

Therefore, the task of analyzing the current state and requirements for the architecture of modern information systems in the field of telecommunications, and the study of ways to transform it into the architecture of the information ecosystem of the digital service provider (DSP) is relevant.

Analysis of recent research and publications

Current state and trends in the field of telecommunications.

In the field of technology, the main trend is the transition to digital IMS architecture, and the introduction of LTE/5G, IoT/IoE services [1].

The IMS network provides all telecommunication services (voice, data, messaging, and multimedia) based on packet-switched IP networks - in contrast to "classic" networks, where voice switching was used for voice transmission. IMS architecture is a set of network functions, divided into three levels - transport, control and application - and connected by standardized interfaces. The network function is not tied to the specific technical implementation of the operator's core network.

The components of the IMS architecture are implemented as VNF - virtual network functions. This allows, in particular, to implement the concept of Network as a Service (NaaS), when the network infrastructure provides its resources as abstract network services through a technology-neutral API (e.g. TMF Open API - TMF909 NaaS API Component Suite). At the same time, the specifics of the service implementation are hidden from higher-level systems - BSS/OSS, which allows you to change the network infrastructure and technologies without affecting the BSS/OSS level system.

Thus, the IMS architecture allows you to build a software-defined network (SDN), which blurs the line between network and information infrastructures - there is a convergence of network and IT components into a single whole. As a result, the role of OSS/BSS (OSS in the first place) as a system that provides not only automation of operational and business processes, but also management

and orchestration of network services, and becomes one of the main assets and resources of the operator.

We can also note the rapid development of M2M/IoT/IoE services and technologies and multimedia technologies that require appropriate communication and IT infrastructure for their implementation.

Development of related information technologies, such as big data, artificial intelligence (AI), and cloud computing, etc. potentially should give new qualities to OSS/BSS.

Standardization.

Given the complexity of business processes and organizational structure of the industry, and as a result - the diversity of information systems that automate these processes, and the difficulty of their integration, an important task is to standardize processes and related systems.

The first efforts to standardize information systems in the field of telecommunications and requirements for them were made in 1982 by the international organizations ISO (International Organization for Standardization) and ITU (International Telecommunication Union), which issued relevant standards (M.3000 series) and proposed a pyramid model OSS/BSS - Telecommunications Management Network (TMN) model.

The next step was to develop a number of standards for the non-governmental international organization TeleManagement Forum (TM Forum), established in 1995. A multilevel business process model - Business Process Framework (eTOM) [2], (fig. 1), a multilevel application model - Application Framework (TAM) [3] and a generalized information model - Information Framework (SID) [4]. These models later became part of the NGOSS (Next Generation Operations Systems and Software) Open Architecture Model, and later the ODA.

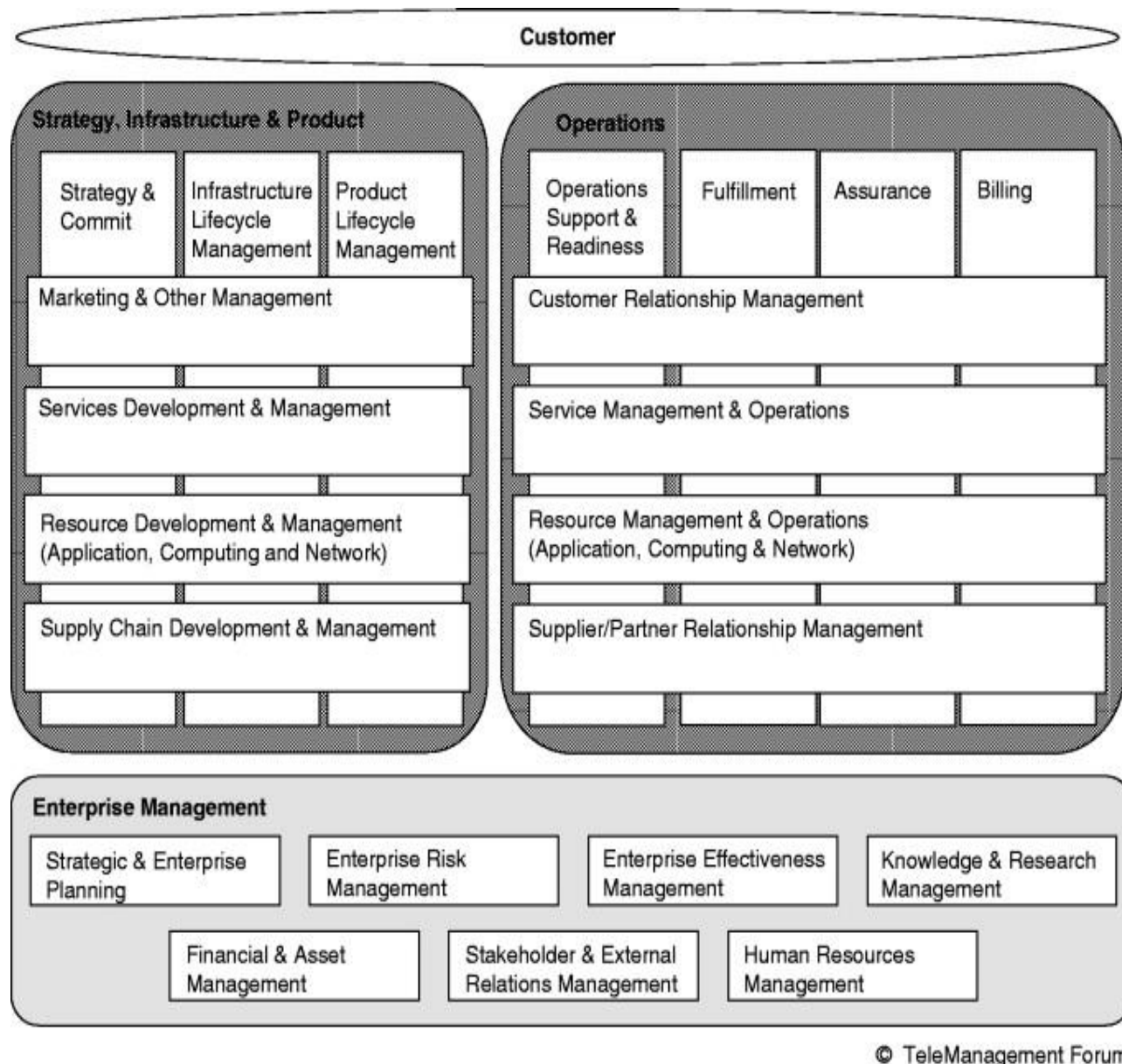


Fig. 1. Multilevel model of eTOM business processes

Recent initiatives of the TM Forum are related to the development of standards for Open Digital Architecture (ODA) [5] and open API (Open API), as part of it.

The telecommunications market is characterized by high competition, the need for large investments in infrastructure, relatively low profitability, reduced ARPU. This requires telecom operators and providers to change the business model.

An important task for telecom operators is the transition from the role of communications service provider (CSP), which provides only transport for more profitable businesses (content providers, information and communication services, digital advertising, IoT services, etc.) to the role of digital services provider (DSP) [6], which provides a full range of digital services based on existing network infrastructure and partner information ecosystem.

In these conditions, one of the key competitive advantages is the reduction of time to launch new services and products (time-to-market), which is determined by the readiness of network infrastructure to new technologies and OSS / BSS readiness to support new products and business processes. Fierce competition on the one hand, and new technologies on the other contribute to the spread of new business models (OTT, B2B2x / B2G2x), when several partner operators are involved in providing services to the end user.

Customer expectations, at the present stage of development of technologies and telecommunications market, in addition to traditional requirements for the quality of services, their availability, security, broadband access, also affect aspects of user/operator interaction: the ability to interact and manage services exclusively through digital channels (digital experience), client-oriented and integration of interaction channels (customer-centric and omni-channel experiences). Subscribers are waiting for the possibility of using high-performance real-time services, which can be customized by performance, set of products and service quality (SLA), on-demand services, etc.

Current state of BSS/OSS in the field of telecommunications. Studies [7, 8, 9] demonstrate that the existing information infrastructure of telecom operators and its components, in particular BSS/OSS, do not meet the development of technology, modern market requirements and customer expectations.

Operators with a "history" a significant part of the infrastructure consists of legacy systems, architecture, and often the functionality of which does not meet modern requirements. The information infrastructure of a large operator can number hundreds and thousands of systems, whose functions are partially duplicated, which cannot be fully integrated and have no single management. As a result, complex business processes are implemented as separate fragments that require manual coordination and orchestration.

Much of the information systems used in telecommunications are based on proprietary and/or

closed standards and interfaces. First of all, it concerns information systems – BSS/OSS. Network components (both hardware solutions and software) are usually built on the basis of open standards (RFC, ITU-T, ETSI), which allows the use of equipment from different manufacturers in the network, simplifies integration and reduces operator dependence on suppliers.

In contrast, BSS/OSS solutions are often closed, providing proprietary interfaces. Historically, many large vendors have been focused on providing a complete stack of BSS/OSS solutions, so their solutions support integration with systems of the same manufacturer, but have poor compatibility (by interfaces, architecture and data models) with systems of other vendors.

Leading providers (vendors) of OSS/BSS are trying to follow the standards and modern requirements for architecture; currently the compliance of the architecture and functionality of OSS/BSS with reference models, (fig.2), at least its declaration, is a prerequisite for success in the information systems market for telecommunications. In practice, such compliance is not always complete, and vendors are forced to maintain inherited systems and solutions, the development of which has a long history and the architecture of which does not meet modern requirements.

Telecom operators usually follow one of two strategies when implementing BSS/OSS. The "best-of-breed" approach provides a choice of solutions of different vendors, each of which best takes into account the requirement of the operator and the specifics of his business processes. The bottleneck here is the need to integrate heterogeneous systems, and integration costs are often comparable to the cost of BSS/OSS itself.

The "best-of-suite" approach involves the purchase of basic information systems from one manufacturer. This approach simplifies the integration of information infrastructure components, but requires significant efforts to customize them, and increases the dependence of the operator from vendor.

In the current environment, when applying any of these approaches, both when replacing existing solutions and implementing new systems, a lot of effort is required to customize them, adapt to the requirements of a particular CSP, integration with existing infrastructure and more. It also takes a lot of effort and time to implement new products and business models. The monolithic or tight coupled architecture of such systems means that their modernization (which is usually performed by the supplier according to the CR - change request procedure) requires significant resources and time. There is a high risk that BSS / OSS implementation projects based on outdated architectural principles will end in excess of cost and delay.

If the BSS/OSS component architecture does not support virtual deployment, it requires the operator to provide additional support costs for technical infrastructure (physical servers, data warehouses, etc.), complicates their administration, and scaling.

eTOM Business Process Framework v.14.0

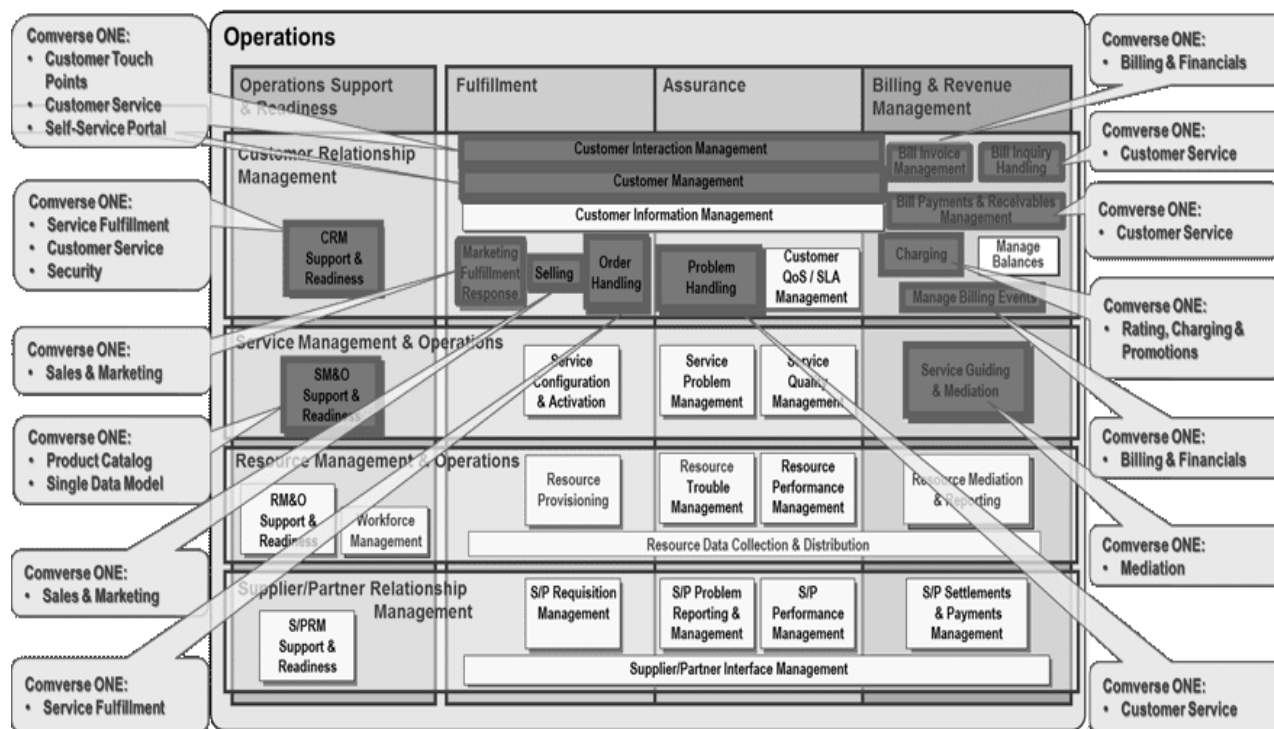


Fig. 2. Projection of Comvers BSS on the model of business processes of eTOM

Task of the study. Thus, it can be concluded that the existing BSS/OSS in the field of telecommunications, in particular their architecture, do not meet modern technological and business challenges, and need to be modernized. Actual is the task of analyzing the requirements for the architecture of such systems, and recommendations for the process of transforming the information infrastructure of the communication operator into an open ecosystem that would allow the introduction of new business models and products, increase the competitiveness and loyalty of subscribers and partners.

The **purpose of this article** is to analyze high-level business and system requirements for information systems in the field of telecommunications, to formulate requirements for the architecture of such systems, and to provide recommendations for the process of transformation of the existing information architecture of telecommunications operators into an open digital architecture.

Materials and methods used in the study - system and business analysis [10, 11], methods of analysis of software requirements, etc.

Study results

BSS/OSS requirements analysis. Briefly, the requirements for BSS/OSS can be formulated as follows: they must meet the modern level of technology development and business requirements.

Analysis of trends in the development of technologies and market requirements in the field of

telecommunications, which were discussed above, allows to formulate the following functional and non-functional requirements for BSS/OSS information systems in the field of telecommunications:

- End-to-end automation of all business processes, which eliminates or minimizes staff interference not only in routine operations, but also in complex processes such as introducing a new product or connecting a new partner, and minimizes the need for manual orchestration of processes.
- Reduction of operating (OPEX) and capital (CAPEX) costs for information systems deployment, upgrade and maintenance of information infrastructure.
- Reduce the time to launch new services and launch new products and marketing offers (time-to-market).
- Rapid adaptation to changes in technology and business models, as well as the load on the network and / or a separate information system.
- Support for modern communication technologies, including LTE / 5G, in accordance with 3GPP, ITU-T and other standards and requirements.
- Future proof - possibility to implement future communication technologies, services and business models without the need to replace or deeply upgrade key BSS / OSS components and systems.
- On-demand scalability, adaptation to changes in traffic volumes and loads, which should ensure the availability of the systems and services for subscribers in a wide range of loads on the network and information systems of the operator.
- Independence from BSS / OSS vendors, the ability to deploy new components from different vendors and

integrate them into the existing information infrastructure, and / or replace individual infrastructure components.

- Support for the software-as-a-service (SaaS) model, which is necessary in particular for the formation of the operator's partner ecosystem. Possibility of seamless integration with information systems of partners.

- Omnichannel - integration of different channels and seamless switching of communication with the subscriber to the optimal channel with preservation of context, and providing digital experience of subscribers.

Requirements for BSS/OSS architecture.

In order for BSS/OSS to meet the above requirements, their architecture must meet the following principles [8, 9, 12, 13]:

Compliance with open standards is primarily the standards of 3GPP, IETF, ETSI, etc. in part of network functions and interfaces, and TMF ODA in part of software architectures and interfaces. It is open standards that are supported by influential international organizations and constantly evolving, unlike proprietary ones, that ensure that the system will meet the current and future requirements of technologies and the market.

The use of open and standardized interfaces should ensure the integration and interchangeability of information infrastructure elements, its flexibility, and the ability to build an information ecosystem of a communication operator and integration with partners. In the field of telecommunications, this is primarily a set of specifications of the TM Forum Open API.

Component based loose coupled architecture, in contrast to the monolithic architecture of legacy systems, should provide BSS/OSS flexibility, and the ability to adapt them quickly to market changes and technology. There are no single requirements for using any single approach: it can be a microservice architecture [12], containerization, SOA, etc. Built on such principles, the system is a network of services that interact with each other and with external systems through the API; management and orchestration of services is performed in accordance with business policies and process models.

Cloud native architecture [9] - the readiness of the solution to complete virtualization and to deploy in an NFV environment, in a private or public (Amazon Web Services, Google Cloud, Microsoft Azure) cloud. This opens the way to scalability and, accordingly, availability of solutions for loads varying in a wide range, as well as the ability to provide software as a service - SaaS.

Model driven and configuration driven architecture. This means that changing processes in the system does not require modification at the level of program code (customization), but is performed by configuring the existing components of the system and integrating new ones. This approach requires a global product and service, a registry of network resources, and the use of modeling tools and standards such as TOSCA to describe services, YANG to describe resources, and BPMN 2.0 to model processes [14].

Support for CI/CD (Continuous Integration and Deployment) processes should ensure the ability to

quickly and safely update components and (sub) systems that are already in operation.

It should be noted that the specified requirements are closely related to each other. For example, standard compliance includes standardized APIs and use of simulation standards, component architecture related to APIs and virtualization, etc.

Thus, the goal is to transform the outdated monolithic information architecture of communication operators into an open information ecosystem built on modern architectural principles and standards, (fig. 3). This should ensure that telecom operators turn into DSP, that provides service on the basis of communication network and information infrastructure. The target architecture should be built on the basis of loose coupled components that are deployed in the public or private cloud and are connected to each other and to external systems using an open standardized API [15].

A significant factor affecting the process of architecture transformation is the inability to simultaneously replace all legacy systems and components of existing information infrastructure. This is due to the large amount of necessary capital investments and high risks of untimely completion of such large-scale projects.

In the process of transforming the BSS/OSS architecture, several tasks must be solved:

- Decomposition of existing information systems into set of loose coupled components according to the model of processes and systems (eTOM and TAM, Open Digital Architecture), and transfer of components to the cloud environment. Legacy systems that cannot be transformed in this way must be replaced with new ones.

- Integration of product catalogs. This requires the analysis and revision of existing products in order to unify and simplify them, and the development of a single data model that allows you to describe all existing products, provide the possibility of their configuration and will meet current standards (including SID). Data from all available legacy directories must be combined into a single shared product catalog and accessible to all information components (usually via the API). At the transition stage, federalization or data consolidation can be used.

- Implementing a single environment for modeling, performing and orchestrating of business processes. This requires analysis and revision of existing processes for the purpose of their unification and digitalization, development of a common model using modern tools of modeling business processes (for example, BPML/BPME).

- Integration of channels of interaction with subscribers and partners, and transition to omnichannelity. This requires the separation of functions that provide interaction with subscribers and partners, from systems that implement basic processes and services, to a separate integration level that will ensure their orchestration.

- The transition of network infrastructure to IMS/SDN architecture, which meets modern standards, is able to support modern telecommunications technologies such as LTE/ 5G, Edge, Io T/Io E and provides real-time management of network resources by OSS/BSS

information systems, including allocation of resources according to the current needs of the subscriber.

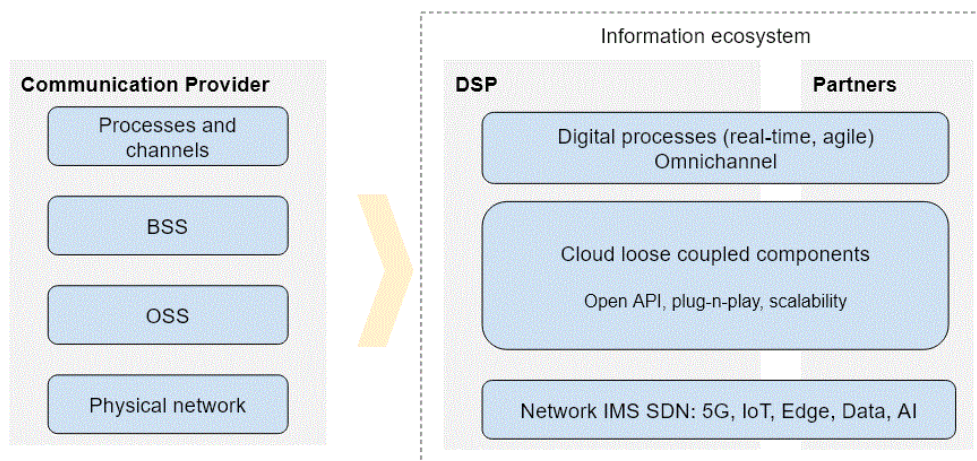


Fig. 3. BSS/OSS architecture transformation

In the process of transformation of the information architecture of telecom operators it is possible to apply several strategies [16]:

- greenfield – parallel deployment of a new information infrastructure, built on the declared principles, and gradual transfer of individual products and/or services and/or business segments. For example, first, prepaid mobile services and subscribers are transferred to the new infrastructure, then a contract mobile, then a fixed one. Thus, at the transition stage, there are two infrastructures - old and new - between which processes and services are distributed.

- brownfield – gradual replacement of individual legacy systems with new ones that meet the above requirements for architecture. The outdated system is simultaneously decommissioned; its functions are transferred to new components. For example, first, CRM is replaced, then systems of interaction with partners, billing system and so on. It should be understood that the improvement and modernization of existing legacy systems in such a way that they meet the above requirements for architecture are possible only from their developers and vendors. Systems that are no longer developed and unsupported must be unopposed.

- carve-out – gradual deployment of new infrastructure and transfer to it certain functions of existing systems. The inherited system, all functions of which have moved to the new infrastructure, is decommissioned.

Each of the above approaches has its own characteristics, advantages and disadvantages. The strategy should be chosen taking into account the existing architecture, business processes, the available project budget, etc. For large telecom operators, the parallel deployment of the new infrastructure in full requires significant resource costs – budget and time – until the first results are received. Therefore, a combined (brownfield and/or curve-out) strategy is usually used for gradually transition to a new architecture.

It is proposed to use intermediate (migration) architecture in the process of transformation and integration of all existing components and systems using a unified API, (fig. 4). It is proposed to use TMF Open API

as a unified interface, which is the actual standard and already used by many telecommunication operators and vendors [17]. The presence of an open standardized interface will allow you to begin the transition to omnichannel interaction with subscribers, unify integration with partners and begin the development of an affiliate ecosystem.

The target architecture must be built using a multi-level component approach [18]. This will transform the monolithic architecture into a poorly connected one, and will create the possibility of gradual modernization as well as reuse of the component. New components must meet the above architectural requirements, in particular, be placed in a virtual (cloud) environment and integrated with Open API.

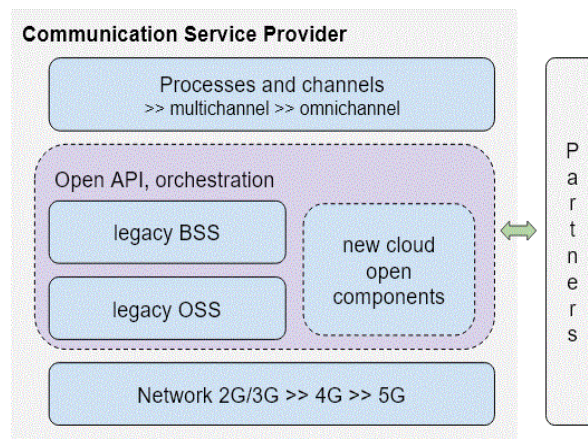


Fig. 4. BSS/OSS intermediate architecture

In parallel, the final transition of the communication network to the IMS/SDN architecture should take place, which will allow implementing the latest services and ensuring the convergence of network and IT-component into a single system.

Conclusions

The transition to the proposed principles of OSS/BSS architecture should ensure compliance of such systems with market requirements and new technologies,

adapt to their changes, allow operators to increase competitiveness, reduce operating and capital costs, etc. The aim should be to create technological conditions for the development of an open digital ecosystem and the transformation of the telecom operator into a digital service provider.

The process of transformation of the architecture of the information infrastructure of telecom operators on the

basis of an intermediate architecture built on the basis of the open standardized Open API TMF is proposed. This process should ensure a gradual transition to the target architecture.

The direction of further research is to improve the process of architecture transformation, detailing and analysis of intermediate architecture options.

References

1. ElNashar, A., El-saidny, M. A. (2018), *Practical Guide to LTE-A, VoLTE and IoT: Paving the way towards 5G*. WILEY, 480 p.
2. Yakunin, K., Kalimoldayev, M., Mukhamediev, R.I., Mussabayev, R., Barakhnin, V., Kuchin, Y., Murzakhmetov, S., Buldybayev, T., Zhumabayev, A., Ospanova, U., Yelis, M., Gopejenko, V., Meirambekkyzy, Z., Abdurazakov, A. (2021), "KazNewsDataset: Single Country Overall Digital Mass Media Publication Corpus", Data 2021, 6, 31. DOI: <https://doi.org/10.3390/data6030031>, <https://www.mdpi.com/journal/data>, Licensee MDPI, Basel, Switzerland.
3. TMForum (2021), "Business Process Framework (eTOM)", available at: <https://www.tmforum.org/business-process-framework> (last accessed 15.12.2021).
4. TMForum (2021), "Application Framework (TAM)", available at: <https://www.tmforum.org/application-framework> (last accessed 15.12.2021).
5. TMForum (2021), "Information Framework (SID)", available at: <https://www.tmforum.org/information-framework-sid>. (last accessed 15.12.2021).
6. TMForum (2021), "Open Digital Architecture (ODA)", available at: <https://www.tmforum.org/oda/> (last accessed 15.12.2021).
7. Goel, S. (2020), "The journey from CSP to DSP. Informa PLC", available at: <https://telecoms.com/opinion/the-journey-from-csp-to-dsp/> (last accessed 15.12.2021).
8. Torosyan, E. K., Shekhovtsova, E. I., Akopyan, A. A. (2019), "Analysis of end-to-end methods for the basic architecture of telecom operator systems automation" ["Analiz metodov skvoznoy avtomatizatsii bazovoy arkhitektury sistem operatora svyazi"], *Petersburg economic journal*, No. 2., P. 172–177.
9. Newman, M. (2020), "FUTURE BSS. Say goodbye to software customization", available at: <https://inform.tmforum.org/research-reports/future-bss-say-goodbye-to-software-customization/> (last accessed 15.12.2021).
10. TMForum (2020), "Cloud native OSS/BSS", available at: <https://inform.tmforum.org/research-reports/cloud-native-oss-bss/> (last accessed 15.12.2021).
11. IIDA (2015), "A Guide to the Business Analysis Body of Knowledge (BABOK Guide)", IIBA, P. 512.
12. Wiegers, K. I., Bitti, J. (2019), "Software Requirements (Developer Best Practices) 3rd Edition" ["Razrabotka trebovaniy k programmnomu obespecheniyu. 3-e izdanie, dopolnennoe"], BHV, Spb., P. 736.
13. Mortensen, M. H., Offredo-Zreik, L. (2020), "Microservices-Based Cloud Native Modernization of OSS/BSS with Open Source. Informa PLC", available at: <https://telecoms.com/intelligence/microservices-based-cloud-native-modernization-of-oss-bss-with-open-source/> (last accessed 15.12.2021).
14. McElligott, T. (2019), "Future OSS. Toward an Open Digital Architecture. TMForum", available at: <https://inform.tmforum.org/research-reports/future-oss-towards-an-open-digital-architecture/> (last accessed 15.12.2021).
15. Smidovych, L.S., Kalmykov, A.V. (2010), "Perspective directions for the development of business processes automation technologies" ["Perspektivnye napravleniya razvitiya tekhnologiy avtomatizatsii biznes-protsessov"], *Radioelectronic and Computer Systems*, Vol. 4 (45), P. 75–81.
16. Turkington, I. (2021), "Open Digital Framework, Open Digital Architecture. ITU-T". *Virtual forum on The Role of Standards in Accelerating Digital Transformation for Cities and Communities*, P. 42.
17. Andersen, L. (2021), "Designing lasting simplification into a BSS contract. TMForum", available at: <https://inform.tmforum.org/digital-transformation-and-maturity/2021/08/designing-lasting-simplification-into-a-bss-contract/> (last accessed 15.12.2021).
18. Newman, M. (2021), "Next generation BSS: the Telefónica way. TM Forum", available from: <https://inform.tmforum.org/research-reports/next-generation-bss-operator-and-vendor-perspectives> (last accessed 15.12.2021).
19. Fedorovich, O.Ye., Babich, A.V. (2016), "Optimization of multilevel component architecture software system" ["Optimizatsiya mnogourovnevnoy komponentnoy arkhitektury programmnoy sistemy"], *Information Processing Systems*, 2016. Vol. 1 (75), P. 175–177.

Received 14.01.2022

Відомості про авторів / Сведения об авторах / About the Authors

Смідович Леонід Сергійович – кандидат технічних наук, доцент, Національний аерокосмічний університет ім. М. Є. Жуковського "ХАІ", доцент кафедри Комп'ютерних наук та інформаційних технологій; м. Харків, Україна; e-mail: lsonlinels@gmail.com; ORCID: <https://orcid.org/0000-0001-6156-9506>.

Смидович Леонид Сергеевич – кандидат технических наук, доцент, Национальный аэрокосмический университет им. Н. Е. Жуковского "ХАИ", доцент кафедры Компьютерных наук и информационных технологий; г. Харьков, Украина.

Smidovych Leonid Serhiyovych – PhD (Engineering Sciences), Associate Professor, National Aerospace University named after N. E. Zhukovsky "Kharkiv Aviation Institute", Associate Professor at the Department of Computer Science and Information Technologies; Kharkiv, Ukraine.

Давидовський Юрій Костянтинівич – аспірант кафедри комп'ютерних наук та інформаційних технологій, Національний аерокосмічний університет ім. М.Є. Жуковського "ХАІ", м. Харків, Україна; e-mail: davidovskiy2350@gmail.com; ORCID: <https://orcid.org/0000-0003-2813-4169>.

Давыдовський Юрій Константинович – аспірант кафедри комп'ютерних наук і інформаційних технологій, Національний аерокосмічний університет ім. Н.Е. Жуковського "ХАІ"; г. Харків, Україна.

Davydovskiy Yurii – graduate student of the Department of Computer science and information technologies, National Aerospace N.E. Zhukovskiy University "Kharkiv Aviation Institute "; Kharkov, Ukraine.

ПРОЦЕСИ ТРАНСФОРМАЦІЇ ІНФОРМАЦІЙНОЇ АРХІТЕКТУРИ ОПЕРАТОРА ЗВ'ЯЗКУ

Предметом дослідження статті є архітектура інформаційних систем операторів зв'язку (BSS/OSS), вимоги до неї та процеси її трансформації. **Метою** є трансформація наявного інформаційного середовища оператора зв'язку у відкриту цифрову екосистему, відповідно до бізнесових та технологічних вимог до функціоналу та архітектури інформаційних систем операторів зв'язку, що дозволить впровадити нові бізнес-моделі, послуги та продукти, канали комунікації із абонентами тощо, та підвищити конкурентоздатність. В роботі вирішуються наступні **задачі**: аналіз сучасного стану сфери телекомунікацій та інформаційних систем, що використовуються операторами зв'язку; аналіз вимог, що висувуються до інформаційних систем підтримки бізнесу та операційної діяльності, збір та аналіз вимог до архітектури сучасних інформаційних систем в сфері телекомунікацій, аналіз процесу трансформації наявної архітектури BSS/OSS оператора зв'язку у цільову архітектуру інформаційної екосистеми цифрового оператора комунікацій, розробка рекомендації щодо проміжної архітектури. Для вирішення зазначених завдань були використані **методи** системного та бізнес аналізу, зокрема методи аналізу вимог до ПЗ, технології розробки архітектури програмних додатків. Отримані наступні **результати**. Проаналізовано сучасний стан галузі телекомунікаційних послуг, особливості наявних інформаційних систем операторів зв'язку (BSS/OSS) та вимоги які висувуються до архітектури таких систем. Розглянуто та проаналізовано стратегії переходу до сучасної інформаційної архітектури. Розглянуто цільову архітектуру інформаційної екосистеми провайдера цифрових послуг та процес трансформації наявної архітектури BSS/OSS у цільову архітектуру. Запропоновано проміжну міграційну архітектуру. **Висновки**: перехід до сучасної інформаційної архітектури забезпечить відповідність систем BSS/OSS вимогам ринку та нових технологій, адаптацію до їх змін, дозволить операторам підвищити конкурентоздатність, зменшити операційні та капітальні витрати, створить технологічні умови для розбудови відкритої цифрової екосистеми та перетворення оператора зв'язку у провайдера цифрових послуг. Для трансформації рекомендується використовувати проміжну архітектуру на основі відкритого стандартизованого API TMF.

Ключові слова: BSS; OSS; інформаційна архітектура; телекомунікації; оператор зв'язку; аналіз вимог; провайдер цифрових послуг; відкрита інформаційна екосистема; омніканальність; процеси трансформації архітектури.

ПРОЦЕССЫ ТРАНСФОРМАЦИИ ИНФОРМАЦИОННОЙ АРХИТЕКТУРЫ ОПЕРАТОРА СВЯЗИ

Предметом исследования статьи является архитектура информационных систем операторов связи (BSS/OSS), требования к ней и процессы ее трансформации. **Целью** является трансформация существующей информационной среды оператора связи в открытую цифровую экосистему, в соответствии с бизнес-технологическими требованиями к функционалу и архитектуре информационных систем операторов связи, что позволит внедрить новые бизнес-модели, услуги и продукты, каналы коммуникации с абонентами и повысить конкурентоспособность. В работе решаются следующие **задачи**: анализ современного состояния сферы телекоммуникации и информационных систем, используемых операторами связи; анализ требований, предъявляемых к информационным системам поддержки бизнеса и операционной деятельности, сбор и анализ требований к архитектуре современных информационных систем в сфере телекоммуникаций, анализ процесса трансформации существующей архитектуры BSS/OSS оператора связи к целевой архитектуре информационной экосистемы цифрового оператора коммуникаций, разработка рекомендации по промежуточной архитектуре. Для решения указанных задач были использованы **методы** системного и анализа, в частности методы анализа требований к ПО, технологии разработки архитектуры программных приложений. Получены следующие **результаты**. Проанализировано современное состояние отрасли телекоммуникационных услуг, особенности имеющихся информационных систем операторов связи (BSS/OSS) и требования, предъявляемые к архитектуре таких систем. Рассмотрены и проанализированы стратегии перехода к современной информационной архитектуре. Рассмотрена целевая архитектура информационной экосистемы провайдера цифровых услуг и процесс трансформации существующей архитектуры BSS/OSS в целевую архитектуру. Предложена промежуточная миграционная архитектура. **Выводы**: переход к современной информационной архитектуре обеспечит соответствие систем BSS/OSS требованиям рынка и новых технологий, адаптацию к их изменениям, позволит операторам повысить конкурентоспособность, уменьшить операционные и капитальные затраты, создаст технологические условия для построения открытой цифровой экосистемы и преобразования оператора связи в цифрового провайдера услуг. Для трансформации рекомендуется использовать промежуточную архитектуру на базе открытого стандартизированного API TMF.

Ключевые слова: BSS; OSS; информационная архитектура; телекоммуникации; оператор связи; анализ требований; провайдер цифровых услуг; открытая информационная экосистема; омниканальность; процессы трансформации архитектуры.

Бібліографічні описи / Bibliographic descriptions

Смідович Л. С., Давидовський Ю. К. Процеси трансформації інформаційної архітектури оператора зв'язку. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 47–54. DOI: <https://doi.org/10.30837/ITSSI.2022.19.047>

Smidovych, L., Davydovskiy, Y. (2022), "Processes of the telecom operator's information architecture transformation", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 47–54. DOI: <https://doi.org/10.30837/ITSSI.2022.19.047>

W. CAO, V. KOSENKO, S. SEMENOV

STUDY OF THE EFFICIENCY OF THE SOFTWARE SECURITY IMPROVING METHOD AND SUBSTANTIATION OF PRACTICAL RECOMMENDATIONS FOR ITS USE

The **subject** of research in the article is a way for evaluating the effectiveness of the software security improving method. The **aim** of the article – study of the effectiveness of the software security improving method and substantiation of practical recommendations for its use. **Tasks** to be solved: analysis of methods for describing the software security testing process and evaluating its effectiveness, developing a scheme and method for evaluating the effectiveness of a method for improving software security, developing a simulation model for the software security testing process, studying the effectiveness of a method for improving software security, researching and substantiating the reliability of the results obtained, developing practical recommendations for using the method. Applied **methods**: system analysis, project approach, heuristic methods of decision making, process models. The **results** obtained: The analysis of the features of the ways for describing the software security testing process and evaluating its effectiveness showed the possibility of taking into account many factors by using the method of dynamics of averages. A way for evaluating the effectiveness of a method for improving software security has been developed, which differs from the known ones by taking into account the scaling factor of the software development process by introducing security testing specialists. With the help of an improved method, the hypothesis of increasing the efficiency of the security process using the developed method by reducing the relative damage indicator at all stages of the software life cycle, depending on the possible duration of a cyber-intrusion, was proved. The substantiation of the reliability of the results of mathematical modeling has been carried out. A number of practical recommendations on the use of the method of improving software security are given and some shortcomings are highlighted, which allow the conclusion that further research is possible.

Keywords: software safety; efficiency evaluation; reliability of mathematical modeling results; practical recommendations.

Formulation of the problem.

The conducted research of the security testing process [1], as well as the synthesis of the method of automated penetration testing [2], show the impossibility of complete processing of data on SW vulnerabilities and assessing them with 100% accuracy. Therefore, methods for approximate evaluation of the effectiveness of existing testing approaches based on modern approaches to mathematical formalization have become widespread [3, 4]. However, their use also implies the adaptation of the main provisions to changes in the SW testing process, and the reduction of input data uncertainty factors.

Literature analysis.

One of the possible ways to describe the process of SW security testing is the method of dynamics of averages. The advantages of this method are simplicity, the ability to take into account many factors (the availability of active and passive testing tools, the capabilities of penetration testing specialists and DevSecOps, etc.), the availability of analytical solutions [5-9]. Improving the method for evaluating the effectiveness of the developed method for improving SW security based on the method of dynamics of averages is based on the following assumptions. According to the law of large numbers, data on possible SW security threats, as well as means of countering them, are close to average (mathematical expectations), which makes it possible not to consider the details associated with the random state of a single element or SW function, and consider the SW security process as deterministic [5]. With this assumption, all test indicators will also not be random variables - they will be replaced by the corresponding mathematical expectations. The sequence of cyberattacks

is represented as a Poisson flow of events [6, 7]. A technique is also used, which consists in the transition from the stream of cyberattacks to the stream of cyberattacks that have achieved results, which is also considered to be Poisson. A cyber attack is called successful if it implements an existing threat. Another possible way to formalize the SW security process is to use the theory of continuous Markov processes [8]. A process occurring in a system is called Markovian if, for each moment of time, the probability of any state of the system in the future depends only on its state at the present moment and does not depend on how the system came to this state. The flow of cyberattacks and the flow of cyberattacks that achieve results are also considered to be Poisson.

Method for evaluating the effectiveness of the developed method for improving SW security

The block diagram of the method for evaluating the effectiveness of the developed method for improving SW security and the labeled graph of system states is shown in fig. 1.

The method for evaluating the effectiveness of the developed method includes the following steps:

1. Analysis of possible SW threats. Meaningful statement of the research problem.
2. Development and synthesis of the main components of the assessed system "SW - Hacker" into a block diagram.
3. Mathematical formalization of the process of finding the numbers of states in differential form in accordance with the method of average dynamics.
4. Formalization of the input data, as well as additional conditions for solving the problem.
5. Solution of a formalized problem.

6. Fixation and approximation of simulation results.

7. Substitution of results and calculation of performance indicators of the developed method for improving SW security.

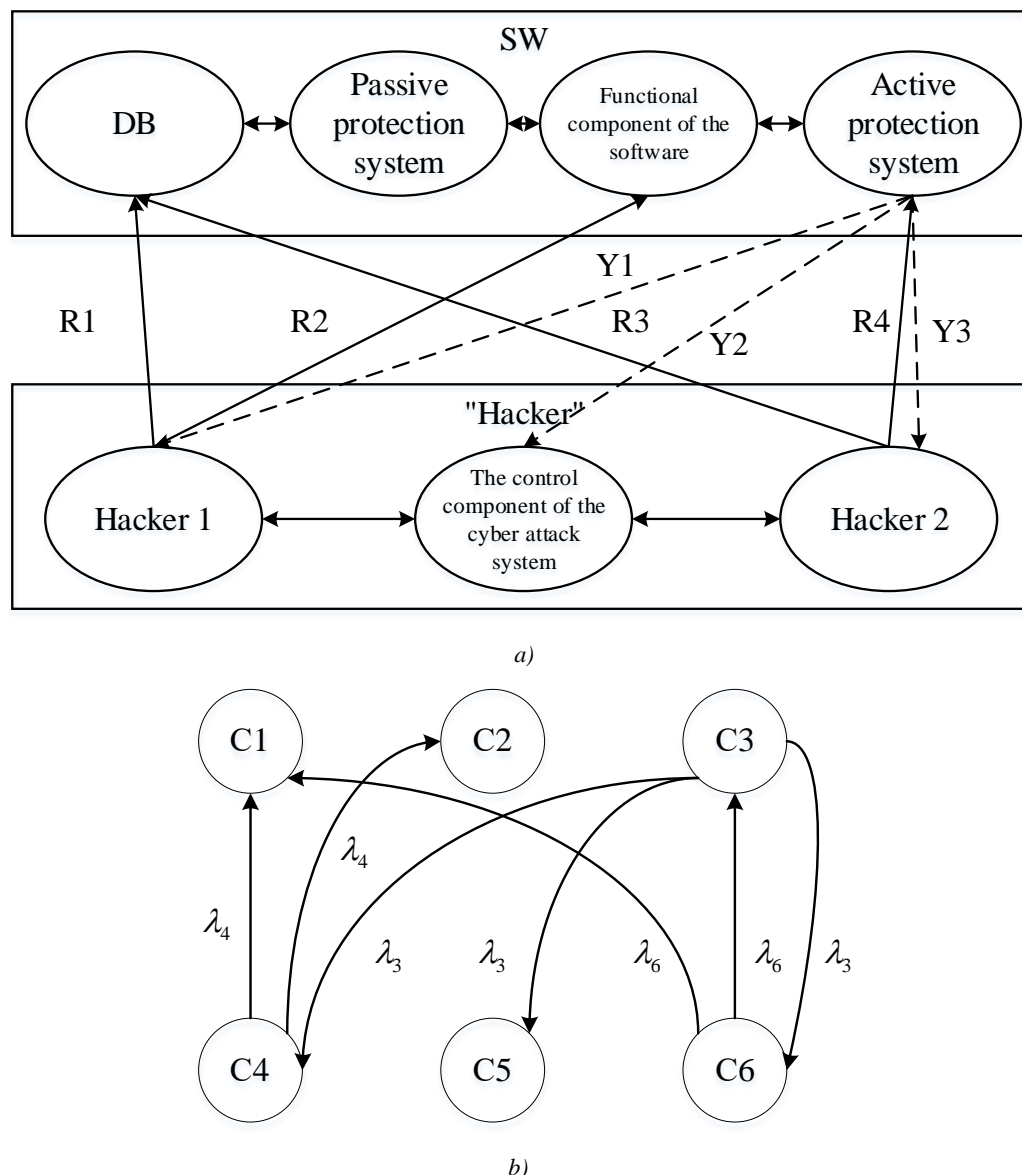


Fig. 1. Structural diagram of the method for evaluating the effectiveness of the developed method for improving SW security and a labeled system state graph

The conducted studies and analysis of a number of scientific sources [10, 11] have shown that the effectiveness of the developed method can be quantitatively determined using the loss index ΔQ_i for each SW element or function q_i . This indicator determines the amount of relative damage caused to the tool SW q_i element as a result of cyber attacks on the SW. At the same time, in accordance with scientific works [9, 10], the indicator of relative damage (losses) can be calculated by the formula:

$$\Delta Q_i(t^*) = \frac{q_i(t_0) - q_i(t^*)}{q_i(t_0)} \cdot 100\%, \quad i = 1 \dots n, \quad (1)$$

where t^* is the current moment of time (the moment of the end of the formalization of the process);

$\Delta Q_i(t^*)$ – relative damage (losses) for an SW element or function q_i at a point in time t^* ;

$q_i(t_0)$ – initial potential of the facility q_i at a point in time t_0 ;

n – the number of phase coordinates (the dimension of the vector q) of the simulated system "SW - Hacker".

8. Analysis and generalization of the simulation results and preparation of practical recommendations on the use of the SW security enhancement method and the strategy for its use in firms.

The program developed in the Python environment made it possible to conduct a series of experiments for given conditions. At the same time, the implemented simulation model of the interaction states of the "SW - Hacker" system made it possible to evaluate the

performance indicators of the developed method for improving SW security.

In this model, in order to increase the reliability of the results obtained, the factors and possibilities of active (with the help of DevSecOps, SecDev, etc.) protection and passive protection (using preventive methods of secure programming and testing to detect and localize unsafe elements and SW functions).

For a formalized representation of the dynamic system "SW - Hacker", we introduce variables q_i , $i=1...6$, by which we mean, respectively, the number of states of the following elements: SW installation components, SW database and active protection means, active attack means of the first type, control system and active attack means of the second type of cyber attacks.

The state of the dynamic system "SW - Hacker" as a whole at each moment of time $t \in [t_0, t_N]$ is characterized by a system of ordinary linear differential equations (2), in which the numbers of states q_i , $i=1...6$ are considered as variables:

$$\left\{ \begin{array}{l} \frac{dq_1}{dt} = \bar{q}_1(t) = -\lambda_4 p_{14} R_1 q_4 - \lambda_6 p_{16} R_3 q_6 + \delta q_1 \gamma_1(q_1, Q_1^*); \\ \frac{dq_2}{dt} = \bar{q}_2(t) = -\lambda_4 p_{24} R_2 q_4 + \delta q_2 \gamma_2(q_2, Q_2^*); \\ \frac{dq_3}{dt} = \bar{q}_3(t) = -\lambda_6 p_{36} R_4 q_6; \\ \frac{dq_4}{dt} = \bar{q}_4(t) = -\lambda_3 p_{43} Y_1 q_3; \\ \frac{dq_5}{dt} = \bar{q}_5(t) = -\lambda_3 p_{53} Y_2 q_3; \\ \frac{dq_6}{dt} = \bar{q}_6(t) = -\lambda_3 p_{63} Y_3 q_3. \end{array} \right. \quad (2.)$$

Here $\delta q_1 \gamma_1(q_1, Q_1^*)$ – mathematically formalized representation of the input of additional resources (cybersecurity specialists) to increase cyber protection q_1 ;

$\delta q_2 \gamma_2(q_2, Q_2^*)$ – mathematically formalized representation of the input of additional resources (cybersecurity specialists) to increase cyber protection q_2 ;

Q_1^* and Q_2^* – boundary values of the numbers of states q_1 and q_2 , accordingly, starting from which a reserve is introduced from the passive protection system;

δq_1 and δq_2 – the intensity of the introduction of reserve funds into the composition q_1 and q_2 ;

$\gamma_1(q_1, Q_1^*)$ and $\gamma_2(q_2, Q_2^*)$ – signal functions determined by the formulas:

$$\gamma_1(q_1, Q_1^*) = \begin{cases} 1, & \text{if } q_1(t) \leq Q_1^*; \\ 0, & \text{if } q_1(t) > Q_1^*. \end{cases}$$

$$\gamma_2(q_2, Q_2^*) = \begin{cases} 1, & \text{if } q_2(t) \leq Q_2^*; \\ 0, & \text{if } q_2(t) > Q_2^*. \end{cases}$$

The rule for solving differential equations allows you to introduce variables W_k , $k=1...8$ to denote the coefficients of differential equations:

$$W_1 = -\lambda_4 p_{14} R_1;$$

$$W_2 = -\lambda_6 p_{16} R_3;$$

$$W_3 = 0;$$

$$W_4 = -\lambda_4 p_{24} R_2;$$

$$W_5 = -\lambda_6 p_{36} R_4;$$

$$W_6 = -\lambda_3 p_{43} Y_1;$$

$$W_7 = -\lambda_3 p_{53} Y_2;$$

$$W_8 = -\lambda_3 p_{63} Y_3,$$

where λ_i , $i=1...6$ – the intensity of the attacks carried out by means of q_i , $i=1...6$ respectively;

$p_{i,j}$ – the probability of penetration (hacking) into the SW or its separate component q_i (for example, a database) as a result of an attack by a mean q_j ;

Y_1 , Y_2 and Y_3 – coefficients characterizing the level of quality work of SW security testers;

R_1 , R_2 and R_3 – coefficients characterizing the degree of training of hackers who overcome SW protection.

It should be noted that the general condition: $Y_1 + Y_2 + Y_3 \leq 1$.

After transformations of the system of equations (2), we obtain:

$$\left\{ \begin{array}{l} \bar{Q}_1(t) = W_1 q_4 + W_2 q_6 + \delta q_1 \gamma_1(q_1, Q_1^*); \\ \bar{Q}_2(t) = W_3 q_6 + \delta q_2 \gamma_2(q_2, Q_2^*); \\ Q_3(t) = W_4 q_6; \\ Q_4(t) = W_5 q_3; \\ \bar{Q}_5(t) = W_6 q_3; \\ \bar{Q}_6(t) = W_7 q_3, \end{array} \right. \quad (3)$$

where $\gamma_k(t)$ is a function that characterizes the possibility of choosing the interval for introducing a reserve Δt_{add} (cybersecurity specialists) to reinforce the security of SW elements and functions $q_1(t)$ and $q_2(t)$.

The ratios $Y_1/Y_2/Y_3 = X_1/X_2/X_3$ determine the strategy of software developers in counteracting cyberattacks of the "Hacker". The ratios $R_1/R_2/R_3 = X_4/X_5/X_6$ determine the strategy of the "Hacker" in the organization of breaking SW.

R_1 – share (%) of the "Hacker" funds q_4 involved in a cyber attack on SW elements and functions. It is necessary to note the obligatory fulfillment of the condition: $R_1 \leq 1$. R_2 – the share (%) of the "Hacker"

funds q_4 involved in a cyber attack on the elements and functions q_2 of the SW (the main functional component). Also R_3 , R_4 are the shares (in %) of the means q_6 of the "Hacker" side participating in the information suppression of the means q_1 and q_3 from the side of the SW developer, respectively. The sums $R_1 + R_2$ and $R_3 + R_4$ must be less than or equal to 1. p_{ij} , $i, j = \overline{1, 6}$ is the probability of failure of the i -th type device by the j -th type device.

Modeling and assessment of relative damage.

When modeling, it is necessary to introduce a number of additional conditions:

1. Replenishment of elements in the state C_1 (adding the active and passive phases of counteracting cyberattacks) is carried out at $q_1(t) \leq 80\% q_1(t_0)$. Replenishment of elements in the state C_2 is performed when $q_2(t) \leq 90\% q_2(t_0)$.

2. Each of the elements, functions or systems q_i , $i = \overline{3, 6}$, has only two states: safe and dangerous.

Table 1. Values of the probability of occurrence of a security error SW or successful countermeasures against the "Hacker"

SW				"Hacker"		
P_{14}	P_{16}	P_{26}	P_{36}	P_{43}	P_{53}	P_{63}
0,1	0,08	0,09	0,03	0,08	0,125	0,06

5. Numerical solution of differential equations (2) is performed in a cycle on the time interval from 0 to 180 c.u., with a step of 0.1 c.u.

6. The limitations of the simulation, in which it is impossible to continue the given program, are the cases

3. The input data of the simulation model are as follows:

$$q_1(t_0) = q_2(t_0) = 100, \quad q_3(t_0) = q_5(t_0) = 100\%,$$

$$q_4(t_0) = q_6(t_0) = 50.$$

$$\lambda_3 = 0,12 \text{ hour}^{-1}, \lambda_4 = 0,15 \text{ c.u.}^{-1}, \lambda_6 = 0,21 \text{ c.u.}^{-1},$$

$$\delta q_1 = 0,25 \text{ c.u.}^{-1}, \delta q_2 = 0,15 \text{ c.u.}^{-1}.$$

4. Taking into account the conducted research and expert assessments of specialists from SW development firms, fixed strategies $\{Y\}$ and $\{R\}$ aspects of the "SW - Hacker" system were adopted:

$$Y_1 = Y_2 = 0,3; \quad Y_3 = 0,25; \quad R_1 = 0,5; \quad R_2 = 0,25;$$

$$R_3 = 0,35; \quad R_4 = 0,4.$$

The values of the probabilities of occurrence of a security error SW or successful countermeasures against the "Hacker" will be presented in accordance with table 1.

when: $q_1 \leq 1$; $q_2 \leq 1$; $q_3 \leq 10\%$; $q_4 \leq 1$; $q_5 \leq 10\%$; $q_6 \leq 1$.

Calculate the relative damage $\Delta Q_i(t^*)$ for all q_i where $i=1, \dots, 6$.

The calculation results are presented in table 2.

Table 2. Relative damage $\Delta Q_i(t^*)$ for all q_i for $t^* = 120$ c.u.

SW			Hacker		
$\Delta Q_1(t^*)$	$\Delta Q_2(t^*)$	$\Delta Q_3(t^*)$	$\Delta Q_4(t^*)$	$\Delta Q_5(t^*)$	$\Delta Q_6(t^*)$
29,57	18,31	11,99	64,57	50,45	40,36

In the considered example, the results of the relative damage assessment indicate the feasibility and effectiveness of the security testing system implemented in the model and, at the same time, the possibility of neutralizing the means of the attacking side (using active and passive defense methods).

The effectiveness of the developed SW safety improvement method can be assessed using the curves of the graphs in fig. 2.

As can be seen from the graphs, the use of the developed SW security enhancement method, taking into account the capabilities of automated penetration testing using deep machine learning technology, reduces the relative damage at all stages of the SW life cycle by up to 6 times, depending on the possible duration of the attack.

Justification of the reliability of the results of mathematical modeling.

To substantiate the reliability of the results obtained in sections 2 and 3, a number of experiments were carried out, in accordance with the conditions:

- the SW development team consists of 7 people, including one DevSecOps and one security tester;
- the main SW development management methodology is SCRUM;
- sprints are divided into weeks and rallies are held daily;
- number of experiments $N^* = 500$.

Based on the results of the experiment, a histogram of the frequency of correct detection of an attack with a higher probability [12] was obtained, which is shown in fig. 3.

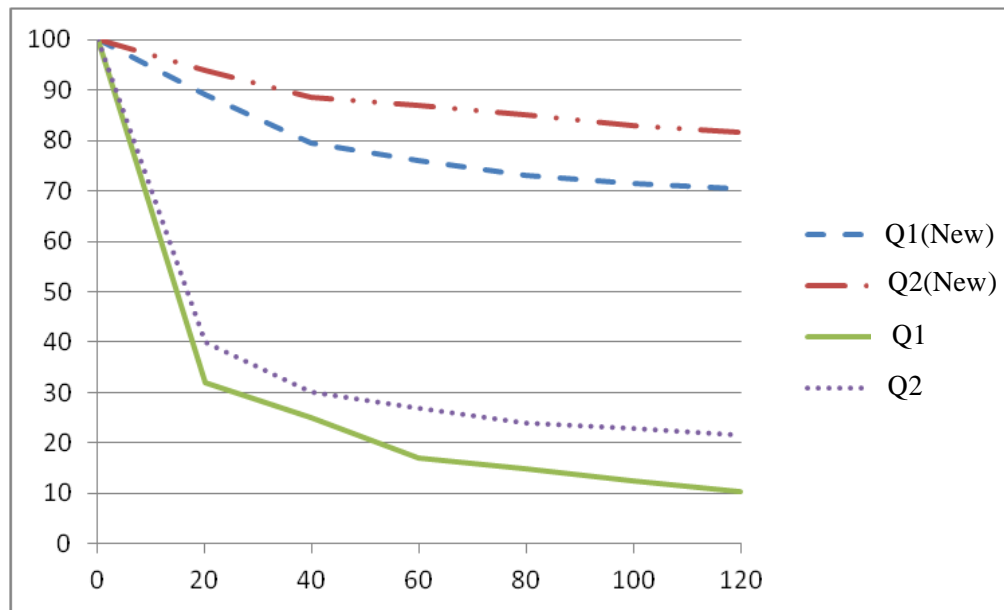


Fig. 2. Graphs of the dynamics of changes $\Delta Q_i(t^*)$ for systems with DevSecOps and a security tester, and also without them

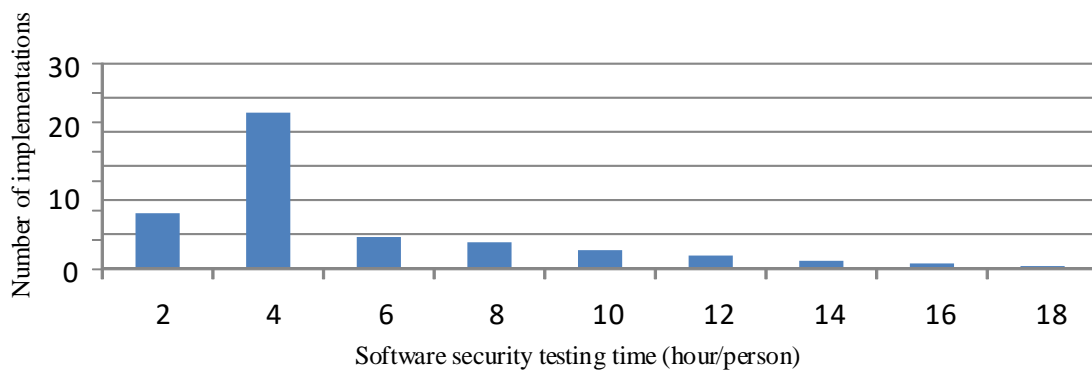


Fig. 3. Histogram of SW security testing time

The hypothesis put forward about the normal distribution of this random variable was tested by Pearson's goodness-of-fit test χ^2 [11]

$$\chi^2 = N^* \sum_{i=1}^k (P_i^* - P_i)^2 / P_i,$$

where k is the number of digits (intervals) of the statistical series;

P_i^* and P_i are "statistical" and theoretical probabilities of "hitting" a given indicator in the i -th category.

The conducted verification proved the plausibility of the hypothesis that the value of the frequency of correct detection of an attack with a higher probability of occurrence is distributed according to the normal law.

Estimates $P_{\text{test}}^{(i)}$ of the mathematical expectation $P_{\text{test}}^{(i)}$ and $\hat{D}_{P_{\text{test}}^{(i)}}$ variance ($\hat{\sigma}_{P_{\text{test}}^{(i)}}$ standard deviation) are obtained, a random variable $P_{\text{test}}^{(i)}$ that characterizes the frequency of correct detection of an attack with a higher probability of occurring:

$$P_{\text{test}}^{(i)} = \frac{\sum_{j=1}^k P_{\text{test}}^{(i,j)}}{N^*}; \quad \hat{D}_{P_{\text{test}}^{(i)}} = \frac{\sum_{j=1}^k (P_{\text{test}}^{(i)} - P_{\text{test}}^{(i,j)})^2}{N^* - 1};$$

$$\hat{\sigma}_{P_{\text{test}}^{(i)}} = \sqrt{\hat{D}_{P_{\text{test}}^{(i)}}}.$$

Using the well-known expression for calculating the confidence probability of the deviation of the relative frequency from the constant probability in independent tests [13], we determine the confidence probability that the value of the characteristic of the frequency of correct attack detection obtained as a result of the experiment with a higher probability of occurrence "does not deviate" from the mathematical expectation $P_{\text{test}}^{(i)}$ by more than 0.05:

$$P\left(\left|\hat{P}_{\text{test}}^{(i)} - P_{\text{test}}^{(i)}\right| < 0.05\right) = 2F\left(0.05 / \hat{\sigma}_{P_{\text{test}}^{(i)}}\right),$$

where F is a Laplace function of the form $F(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ [14].

Also, the results of the experiments showed that for all the studied types of data the confidence probability is that the values of the statistical quantity $P_{test}^{(i)}$ will not deviate from the mathematical expectation $\hat{P}_{test}^{(i)}$ by more than 0.05 and is equal to: $P \approx 0.94$. A comparative study of the results of mathematical modeling and experiment has been carried out. The results of the comparison are shown in fig. 4. in the form of a graph of the probability distribution density of the implementation time of penetration testing algorithms, and the boundaries of the confidence interval corresponding to them:

$$I_{\beta} = \left[\hat{J} - \varepsilon_{\beta}, \hat{J} + \varepsilon_{\beta} \right],$$

in which the true value \bar{J} falls with a confidence probability $\beta = 0,94$ and estimates of its $\hat{t}_{test}^{(i)}$ mathematical expectation.

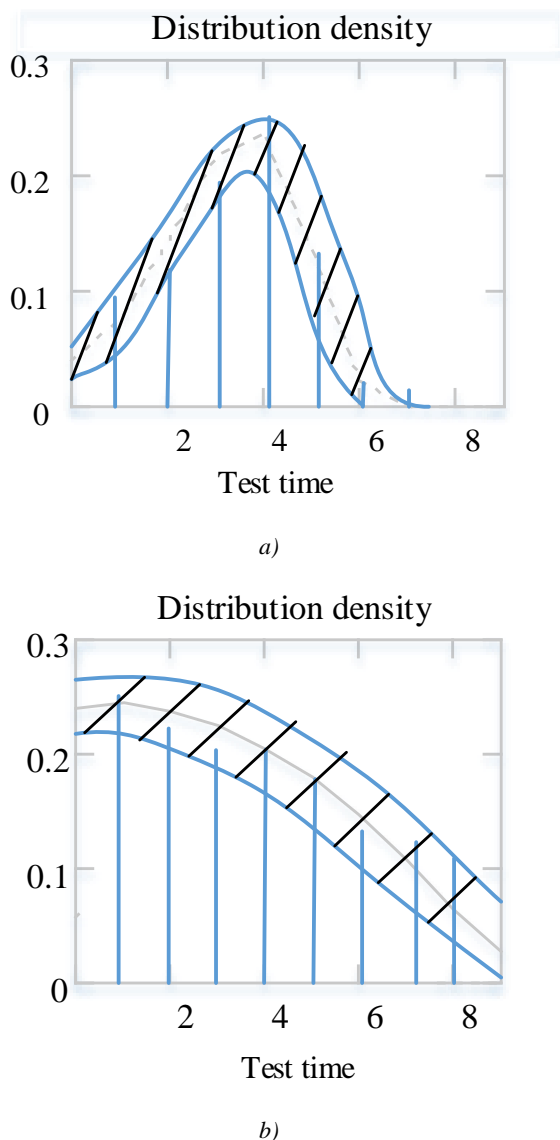


Fig. 4. Graph of the probability distribution density of the implementation time of penetration testing algorithms, the boundaries of the confidence interval corresponding to them, and estimates of its $\hat{t}_{test}^{(i)}$ of mathematical expectation

It can be seen from the graphs that in the key test situation (time $t_{test} \approx 4c.u.$) the "calculated" curve J (solid curve), obtained in accordance with the mathematical model developed in the work, in most practical cases fall into the "averaged" confidence interval (shaded area).

This confirms the reliability of the mathematical model and the analytical expression obtained as a result of mathematical modeling.

Substantiation of practical recommendations on the use of the software security improvement method.

The conducted studies have shown that when designing SW, it is necessary to take into account that the fulfillment of any of the requirements of quality indicators during design may affect other requirements. And in this case, it is necessary to analyze the ratio of benefits and losses for the totality of many quality indicators.

As noted in [13], security is the ability of a system to prevent malicious or accidental actions not provided for in the design, or to prevent disclosure or loss of data. Here we are talking about system security in a general sense. At the same time, improving the security of SW also leads to an increase in the reliability of the system as a whole by reducing the likelihood of successful attacks and their negative impact on the system.

The analysis of the literature [9, 13, 14] and the experience of coding and testing in the development of SW made it possible to identify a number of features of secure programming and implementation of security-related control functions:

- coordination of the SW security policy regarding the technical conditions of the object of implementation (implementation), the formation of restrictions on the use of the product in accordance with the security policy;
- signature-heuristic analysis of source and executable code for potentially dangerous operations and coding errors;
- analysis of security subsystems (password protection subsystem tracing), etc.;
- carrying out all types of manual and automatic testing (functional, stress, load testing and performance testing) taking into account increased security requirements;
- structural analysis of distribution redundancy and integrity control;
- analysis of the presence of covert channels;
- update and modification of SW in accordance with the approved security policy.

A qualitative assessment of these factors and features of secure SW programming allowed us to conclude that an integrated approach to identifying vulnerabilities is expedient and to present common security-related errors of programmers, as well as methods for testing and identifying these errors in the form of a table 3.

Table 3. Code vulnerabilities and SW testing methods

Programming mistakes	Test method	Note
Errors that occur when entering rarely used input data.	Functional testing of rarely used inputs.	Validation of the functionality declared in the document of rarely used input data using boundary or negative testing methods.
Undeclared input parameters and modes associated with an abnormal state of the system and a possible decrease in its performance.	Performance profiling.	Identification of code fragments in which anomalies are observed (unauthorized transfer, transition to an infinite loop, processing of undeclared input data, etc.) performance decrease.
Errors related to functional safety.	Load and stress testing.	Creation of specific conditions for the program to work with a large amount of input data, increased load on the processor, etc.
Vulnerabilities of the password protection subsystem and other subsystems.	Testing of the authentication subsystem, access control and other information security subsystems.	Defining security management modes, tracing the password system, searching for built-in passwords, etc.
Unintentional (intentional) software bookmarks, bombs.	Signature-heuristic analysis.	Identification of software bookmarks, bombs by signatures of potentially dangerous operations or by read event data.
Incorrect coding.	Signature-heuristic analysis.	Identification of software bookmarks by signatures.
Intentional tabs of a "hooligan" nature, "programmer signatures", "Easter eggs".	Signature-heuristic analysis.	Identification of program bookmarks on the basis of illegitimate identifiers of programmers (messages, constants, hidden "hot keys", etc.).
Hidden channels.	Analysis of traffic and memory in an isolated environment.	Monitoring and auditing logs.
Program bookmarks initiated by a hidden transition.	Redundancy control, signature-heuristic analysis.	Identification of software bookmarks on the basis of hidden data transfer.

Conducted studies have shown that the vast majority of software is based on previously developed software components used for work. For example, programs written in C or C++ depend on the runtime libraries that come with the compiler or operating system. Programs written in Python also include a number of standard libraries, components, or other off-the-shelf software solutions. At the same time, the analysis of the literature [12, 13] showed that there are a number of hidden (undeclared) threats.

One of the undeclared threats to using off-the-shelf software solutions is that even if you write flawless code, your application may be vulnerable due to security flaws in one of the components used. For example, the `realpath()` function of the C programming language library returns the canonicalized absolute path for the path given as an argument. To get the canonicalized absolute path, the function expands all symbolic links. However, some implementations of `realpath()` contain a static buffer that overflows when the canonicalized path is longer than `MAXPATHLEN`. Other common C library functions for which some implementations are known to be prone to buffer overflows are `syslog()`, `getpass()`, and the `getopt()` family of calls.

Since many of these problems have been known for some time, many C libraries now contain fixed versions of these functions. For example, the implementations of `libc4` and `libc5` for Linux contain a buffer overflow vulnerability in `realpath()`, but the problem is fixed in `libc-5.4.13`. On the one hand, this creates the prerequisites for eliminating threats and correcting the defect, but, as studies have shown, this problem has not been fully resolved.

Modern operating systems are known to support dynamic link libraries, or shared libraries. In this case, the library code is not statically linked to the executable file, but is searched for in the environment where the program is installed. Therefore, if our development application designed to work with `libc-5.4.13` is installed in an environment that has an older version of `libc5` installed, the program will be subject to a buffer overflow error in the `realpath()` function.

One solution is to statically link the safe libraries with your application. This approach allows you to capture the implementation of the library you are using. However, this approach has the disadvantage of creating a large executable image on disk and in memory. It also means that your application will not be able to take advantage of new libraries that may fix previously

unknown flaws (both security flaws and others). Another solution is to make sure that the input values passed to an external function stay within limits that are known to be safe for all existing implementations of that function.

It should be noted that similar problems can arise with distributed object systems such as DCOM, CORBA, and other compositional models that are linked at run time.

System integrators and administrators can protect systems from vulnerabilities in off-the-shelf software components (such as libraries) by providing wrappers that intercept API calls that are known to misbehave. Such a wrapper implements the original API functions (usually by calling the original component), but performs additional checks to ensure that known vulnerabilities cannot be exploited by an attacker. To be realistic, this approach requires linking executables at runtime. An example of an approach that implements this method for Linux systems is the libsafe library from Avaya Labs [13].

Shells do not require changes to the operating system and work with existing binaries. They fail to protect against unknown security flaws: if there is a vulnerability in the part of the code that is not caught by the shell, the system will still be vulnerable to attacks.

Related to the approach described is the execution of untrusted programs in a controlled environment that is restricted to specific behavior using custom policies. An example of this approach is Sysrtrace. This approach differs from a secure shell in that it does not prevent the exploitation of vulnerabilities, but it can prevent unexpected secondary actions that exploit authors usually try to perform, such as writing files to a secure location or opening network sockets [14].

Sysrtrace is a policy enforcement tool that provides a way to monitor, intercept, and restrict system calls. Sysrtrace acts as a wrapper for executables, directing them to bypass the system call table. It intercepts system calls and, using the Sysrtrace device, passes them through the kernel and processes system calls [14].

Like a secure shell, a controlled environment does not require any source code or changes to the program being controlled. The disadvantage of this approach is that it is very easy to misformulate policies and break the required functionality of the controlled program. Research has shown that it can be a difficult (and often impossible) task for an administrator to create accurate policy descriptions for large and complex programs whose behavior is not fully understood. Accordingly, the solution of the above problems can become a plan for further research.

References

1. Semenov, S., Weilin, C. (2020), "Testing process for penetration into computer systems mathematical model modification", *Advanced Information Systems*, Vol. 4, Issue 3, P. 133–138. DOI: <https://doi.org/10.20998/2522-9052.2020.3.19>
2. Semenov, S., Weilin, C., Zhang, L., & Bulba, S. (2021), "Automated penetration testing method using Deep machine learning technology", *Advanced Information Systems*, Vol. 5, Issue 3, P. 119–127. DOI: <https://doi.org/10.20998/2522-9052.2021.3.16>
3. Farchi, E., Hartman, A., Pinter, S. (2002), "Using a model-based test generator to test for standard conformance", *IBM Systems Journal*, Vol. 41, Issue 1, P. 89–110. DOI: <https://doi.org/10.1147/sj.411.0089>

Among the main security threats and bugs in Python are the following:

- Command injection (command injection).
- SQL injection.
- XML parsing.
- assert statements.
- Time attacks.
- Polluted site-packages or import path.
- Temporary files.
- Using yaml.load.
- Deserialization of canned data.
- Using the Python runtime system.
- Not installing patches for your dependencies.

Detailed information about each of the listed vulnerabilities is presented in the article [15].

Conclusions

The section contains comparative studies and evaluation of the effectiveness of the method for improving software security. For this purpose, the method for evaluating efficiency has been improved in the work. Its distinguishing feature is taking into account the possibility of scaling the software development process by introducing security testing specialists (DevSecOps, SecDev, as well as penetration testers).

The improved method for evaluating the effectiveness of the software security improvement method is based on the method of dynamics of averages.

With the help of an improved method, the expediency of using the developed method for improving software security is proved, taking into account the capabilities of deep reinforcement learning technology. This will reduce the relative damage at all stages of the SW life cycle by up to 6 times, depending on the possible duration of a cyber-intrusion.

The section substantiates the reliability of the results of mathematical modeling. The results of the experiments showed that for all types of data under study, the confidence probability that the value of the statistical value $P_{test}^{(i)}$ will not deviate from the mathematical expectation $\hat{P}_{test}^{(i)}$ by more than 0.05 and is equal to: $P \approx 0.94$.

The section also provides a number of practical recommendations on the use of the software security enhancement method and highlights some shortcomings, which led to the conclusion that further research is possible.

4. Ali H. Doğru, Veli Biçer (2010), "Modern Software Engineering Concepts and Practices: Advanced Approaches", *IGI Global*, P. 506.
5. Shanahan, L., Sen, S. (2011), "Dynamics of stochastic and nearly stochastic two-party competitions", *Physica A: Statistical Mechanics and its Applications*, Vol. 390, Issue 10, P. 1800–1810. DOI: <https://doi.org/10.1016/j.physa.2010.12.041>
6. Tze Leung Lai, Haipeng Xing (2008), *Statistical Models and Methods for Financial Markets*, Springer New York Softcover reprint of hardcover 1st ed., 356 p.
7. Stephen Boyd, Lieven Vandenberghe (2018), "Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares", *Cambridge University Press*. DOI: <https://doi.org/10.1017/9781108583664>
8. Swart, J., Winter, A. (2010), "Markov processes: theory and examples", available at: // <https://www.uni-due.de/~hm0110/Markovprocesses/sw20.pdf>
9. Kosenko, Nataliia & Kadykova, Iryna & Artiukh, Roman. (2017), "Formalizing the problem of a project team building based on the utility theory", *Innovative technologies and scientific solutions for industries*, P. 53–57. DOI: <https://doi.org/10.30837/2522-9818.2017.1.053>
10. Khalife, Kassem, Krikhovetskiy H.Ya., i H.A. Kuchuk. (2017), "Evaluation of the system software security"["Ocinka vrazlivosti sistemnogo programnogo zabezpechennya"], *Management systems, navigation and communication. Collection of scientific*, 6 (46), Poltava: PNTU, P. 141–44.
11. Semenov S. G., Khalife Kassem, Zakharchenko M. M. (2017), "An improved way to scale agile software development", ["Uovershenstvovannyj sposob masshtabirovaniya gibkoj metodologii razrabotki programnogo obespecheniya"], *Bulletin of NTU "KhPI"*, Kharkiv, Vol. 1, No. 1, P. 79–84. DOI: <https://doi.org/10.20998/2522-9052.2017.1.15>
12. Gmurman V.E. (2003), *Theory of Probability and Mathematical Statistics*, [Teoriya veroyatnostej i matematicheskaya statistika], M., Higher school, 479 p.
13. J.D. Meier, David Hill, Alex Homer, Jason Taylor, Prashant Bansode, Lonnie Wall, Rob Boucher Jr., Akshay Bogawat. (2009), "Microsoft's Guide to Application Architecture Design", available at: // <http://ce.sharif.edu/courses/91-92/1/ce474-2/resources/root/App%20Arch%20Guide%202.0.pdf>
14. Robert Seacord (2013), "Secure Coding in C and C++ Addison-Wesley Professional", P. 600.
15. Anthony Shaw (2018), "10 common security gotchas in Python and how to avoid them", *Hakernoon*, available at: // <https://hackernoon.com/10-common-security-gotchas-in-python-and-how-to-avoid-them-e19f8e265e03>

Received 14.01.2022

Відомості про авторів / Сведения об авторах / About the Authors

Цао Вейлін – науковий співробітник, викладач, департамент ІТ інформаційного центру, Нейцзянський педагогічний університет; e-mail: saowl@njtc.edu.cn; ORCID: <http://orcid.org/0000-0001-8230-5235>.

Цао Вэйлинь – научный сотрудник, преподаватель, департамент ИТ информационного центра, Нейцзянский педагогический университет, Neijiang, Китай.

Cao Weilin – Intermediate grade of experimenter, teacher, Department of IT information Centre, Neijiang Normal University, Neijiang, China.

Косенко Віктор Васильович – доктор технічних наук, професор, Національний університет "Полтавська політехніка імені Юрія Кондратюка", професор кафедри автоматики, електроніки та телекомунікацій, Полтава, Україна; kosvict@gmail.com; ORCID ID – <http://orcid.org/0000-0002-4905-8508>.

Косенко Виктор Васильевич – доктор технических наук, профессор, Национальный университет "Полтавская политехника имени Юрия Кондратюка", профессор кафедры автоматики, электроники и телекоммуникаций, Полтава, Украина.

Viktor Kosenko – Doctor of Sciences (Engineering), Professor of Automation, Electronic and Telecommunication Department of National University «Yuri Kondratyuk Poltava Polytechnic, Poltava, Ukraine.

Семенов Сергій Геннадійович – доктор технічних наук, професор, професор кафедри кібербезпеки та інформаційних технологій ХНЕУ ім. С. Кузнеця, Харків, Україна; e-mail: s_semenov@ukr.net; ORCID: <https://orcid.org/0000-0003-4472-9234>.

Семенов Сергей Геннадьевич – доктор технических наук, профессор, профессор кафедры кибербезопасности и информационных технологий ХНЭУ им. С. Кузнеця, Харків, Україна.

Semenov Serhii – Doctor of Sciences (Engineering), Professor, Professor of the Department of cybersecurity and information technology, Simon Kuznets Kharkiv National University of Economics Kharkiv, Ukraine.

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МЕТОДУ ПІДВИЩЕННЯ БЕЗПЕКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ І ОБҐРУНТУВАННЯ ПРАКТИЧНИХ РЕКОМЕНДАЦІЙ З ЙОГО ВИКОРИСТАННЯ

Предметом дослідження у статті є спосіб оцінки ефективності методу підвищення безпеки програмного забезпечення. **Мета** статті – дослідження ефективності методу підвищення безпеки програмного забезпечення та обґрунтування практичних рекомендацій щодо його використання. **Завдання**, що вирішуються: аналіз способів опису процесу тестування безпеки програмного забезпечення та оцінки його ефективності, розробка схеми та способу оцінки ефективності методу підвищення безпеки програмного забезпечення, розробка імітаційної моделі процесу тестування

безпеки програмного забезпечення, дослідження ефективності методу підвищення безпеки програмного забезпечення, дослідження та обґрунтування достовірності отриманих результатів, розробка практичних рекомендацій щодо використання методу. **Методи**, що застосовуються: системний аналіз, проектний підхід, евристичні методи прийняття рішень, процесні моделі. Одержані **результати**: проведений аналіз особливостей способів опису процесу тестування безпеки програмного забезпечення та оцінки його ефективності показав можливість урахування багатьох факторів шляхом використання методу динаміки середніх. Розроблено спосіб оцінки ефективності методу підвищення безпеки програмного забезпечення, що відрізняється від відомих обліком фактора масштабування процесу розробки програмного забезпечення шляхом впровадження спеціалістів тестування безпеки. За допомогою вдосконаленого способу доведено гіпотезу про підвищення ефективності процесу забезпечення безпеки за допомогою розробленого методу шляхом зниження показника відносної шкоди на всіх етапах життєвого циклу програмного забезпечення залежно від можливої тривалості кібервтрутнення. Проведено обґрунтування достовірності результатів математичного моделювання. Наведено ряд практичних рекомендацій щодо використання методу підвищення безпеки програмного забезпечення та виділено деякі недоліки, що дозволило зробити висновок про можливість подальших досліджень.

Ключові слова: безпека програмного забезпечення; оцінка ефективності; достовірність результатів математичного моделювання; практичні рекомендації.

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ МЕТОДА ПОВЫШЕНИЯ БЕЗОПАСНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ОБОСНОВАНИЕ ПРАКТИЧЕСКИХ РЕКОМЕНДАЦИЙ ПО ЕГО ИСПОЛЬЗОВАНИЮ

Предметом исследования в статье является способ оценки эффективности метода повышения безопасности программного обеспечения. **Цель** статьи – исследование эффективности метода повышения безопасности программного обеспечения и обоснование практических рекомендаций по его использованию. Решаемые **задачи**: анализ способов описания процесса тестирования безопасности программного обеспечения и оценки его эффективности, разработка схемы и способа оценки эффективности метода повышения безопасности программного обеспечения, разработка имитационной модели процесса тестирования безопасности программного обеспечения, исследование эффективности метода повышения безопасности программного обеспечения, исследование и обоснование достоверности полученных результатов, разработка практических рекомендаций по использованию метода. Применяемые **методы**: системный анализ, проектный подход, эвристические методы принятия решений, процессные модели. Полученные **результаты**: проведенный анализ особенностей способов описания процесса тестирования безопасности программного обеспечения и оценки его эффективности показал возможность учета многих факторов путем использования метода динамики средних. Разработан способ оценки эффективности метода повышения безопасности программного обеспечения, отличающийся от известных учетом фактора масштабирования процесса разработки программного обеспечения путем внедрения специалистов тестирования безопасности. С помощью усовершенствованного способа доказана гипотеза о повышении эффективности процесса обеспечения безопасности с помощью разработанного метода путем снижения показателя относительного ущерба на всех этапах жизненного цикла программного обеспечения, в зависимости от возможной продолжительности кибервтрутнения. Проведено обоснование достоверности результатов математического моделирования. Приведено ряд практических рекомендаций по использованию метода повышения безопасности программного обеспечения и выделены некоторые недостатки, что позволило сделать вывод о возможности дальнейших исследований.

Ключевые слова: безопасность программного обеспечения; оценка эффективности; достоверность результатов математического моделирования; практические рекомендации.

Бібліографічні описи / Bibliographic descriptions

Цао Вейлін, Косенко В. В., Семенов С. Г. Дослідження ефективності методу підвищення безпеки програмного забезпечення і обґрунтування практичних рекомендацій з його використання. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 55–64. DOI: <https://doi.org/10.30837/ITSSI.2022.19.055>

Cao, Weilin, Kosenko, V., Semenov, S. (2022), "Study of the efficiency of the software security improving method and substantiation of practical recommendations for its use", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 55–64. DOI: <https://doi.org/10.30837/ITSSI.2022.19.055> r

S. ALYOKHINA, I. NEVLIUDOV, YU. ROMASHOV

COMPUTER SIMULATIONS OF CONTROLLABILITY PROCESSES FOR ROBOTIC WHEELED PLATFORMS TAKING INTO ACCOUNT RESTRICTIONS OF JERK MOTIONS

The computer simulations are considered as the required tool to design the suitable autonomous control systems optimal in different senses and especially in providing the restrictions of the jerk motions for the robotic wheeled platforms. The **subject** matter of this research is the development of the theory and methods for computer simulations of the controllability processes of the robotic wheeled platforms. The **goal** of this research is to consider the jerks of the wheeled platforms, and the jerks are reduced to the limitation of the acceleration time derivative of the mass center of the wheeled platform, so that this derivative is considered as the quantitative estimation of the jerks. The incorrectness in the Hadamard's sense for direct defining the jerks by differentiations of the phase coordinates in the case of computer simulations using the numerical methods is discussed. **Tasks** of this research are in developing the generalized approaches for mathematical modelling and computer simulations and in theoretical receiving of the properties inherent for the wheeled platforms and suitable for verification the computer simulations results, as well as in making the calculations to have the quantitative results about the controllability processes for the particular case of the electromechanical four-wheeled platform under the straight motion with the mode of speeding-up from the state of the rest. **Methods** of this research are based on the Lagrange's equations of second kind, as well as on the electromechanical analogies, and on final representing the mathematical models in the form of the system of the first ordered ordinary differential equations with the initial conditions for further numerical solving. The computer simulations are accomplished by using the Scilab free open source software. **Results** of this research are in the proposed suitable way for computing the jerks by the phase coordinates without its differentiations allow excluding the incorrectness in the Hadamard's sense, as well as in representing the controllability processes for the electromechanical wheeled platform, including the results for the velocities, the accelerations and the jerks which are necessary to illustrate the controllability processes for the robotic wheeled platforms. By comparison with the theoretically established inherent properties of the wheeled platforms it is shown the correctness of the results of the computer simulations. **Conclusions** about this research are that the developed approaches for computer simulations of the controllability processes for the robotic wheeled platforms allow considering influence of the control on the different characteristics including the velocity, the acceleration, as well as the jerk motions which are required for designing the controls optimal in different senses.

Keywords: Robotics; Wheeled Platforms; Jerk; Controllability; Mathematical Model; Computer Simulations.

Introduction

Implementation of the robotic wheeled platform for excluding human from the executed processes is one of the main modern trends of the technological development, and the robotic wheeled platforms are widely used now for industrial, transportation, military, police, house-holding and different other purposes. Developing of the improved control is significantly required to have the necessary properties providing the suitable operating for the robotic wheeled platforms relevant to its operation purposes. It is clearly understood that developing and designing the improved control is mostly based on using the opportunities of improved mathematical modelling and computer simulations of the controllability processes for the wheeled platforms, and this research is exactly directed to developing the approaches for computer simulations of the controllability processes of the robotic wheeled platforms which is critically required for designing the control systems optimal in different senses and relevant for the operation purposes. Thus, the theme of this research is of current interest, because of full agreement with the modern global trend in technological development by means wide implementing the robotic systems including the wheeled platforms for making the different operations supporting by the relevant control systems properties.

Analysis of last achievements and publications

The existed wide interest to the problems about control of the robotic wheeled platforms are due to the wide range of the operation purposes for these platforms, and it can be illustrated by the follows researches [1-3]. Such interest to the control problems is due to the different circumstances [4-6]. First of all, the structural and algorithmic designs of the control systems are principally predefined by the operation purposes of the controlled systems, so wide range of the operational purposes for the robotics wheeled platforms inherent for the modern industries will naturally require the different kinds of the control systems that must be designed. Besides, trends in implementing the robotic systems including based on the wheeled platform cover permanently the new areas of purposes, and it will require developing the new kinds of the control systems for the robotic wheeled platform relevantly to theirs new purposes. The understandable differences between controls required for the robotic wheeled platforms used in the industrial technological processes and for the police purposes is the illustration in necessities in developing the different kinds of the control systems for the robotic wheeled platforms relevantly to their purposes. Secondly, the operational environment can be significantly different even for the robotic wheeled platforms with the identical purposes, and theirs control systems must be relevant to the particular operational environment, so the wide range of the possible operational environments leads to necessities of different kinds control systems for the

robotic wheeled platforms [7]. Really, the different environments have the different and even unpredictable external influences on the robotic wheeled platforms, and the control systems must provide compensations of these influences to have the exactly required current states of the platform. The understandable differences between the wheeled platform on the soft soil and the wheeled platform on the rigid road is the illustration of influencing the environment on the requirements to the control systems. Thus, the wide range of requirements to the control systems for the robotic wheeled platforms lead the permanent interest to the problems from the areas about the structural and algorithmic designs of such control systems, and each separate research deals with some particular problem from this area like controlling the velocity, the acceleration, the path, the jerks and others.

It is understandable that the mathematical models are used to substantiate the structural and algorithmic designs of the control systems including for the robotic wheeled platforms, and computer simulations give the significant opportunities to have the extended imaginations about the designed systems even without having the physical prototypes. Due to these circumstances, the computer simulations of the controllability processes are widely used for designing the structural and algorithmic structures of the control systems for the robotic wheeled platforms, and the relevant methodologies of such simulations must be developed, so developing of the methodologies for mathematical modelling and computer simulations of the robotic wheeled platforms are in current interest at present [8-10]. Interest to methodologies of mathematical modelling and the computer simulations of the robotic wheeled platforms are permanently due to a lot of existed particularities inherent for the different kinds of the wheeled platforms which must be considered in the mathematical models and must be imagined during the computer simulations. These particularities can include the difficultly predicted interactions between the wheels and the soils and the roads, the aerodynamics effects during different velocities, the different kinds of damping in the driving mechanisms. Besides, the perfect mathematical models and computer simulations are significantly required for providing the improved indirect measurements and for designing the intelligent control systems including for the autonomous wheeled platforms.

Highlight of the earlier unresolved parts of the general problem. Aim of the study

It seems that the problems of the robotic wheeled platforms control are fully researched due to a lot of existed publications [1-3], but the most of the existed publications deals with the separate particularities of the general problem about control of the robotic wheeled platforms [4-6], so this general problem is actually not fully researched now. Besides, permanent implementing the new kinds of the robotic wheeled platforms requires the developing the relevant control systems [7], so it seems that the problem about control of the robotic wheeled platforms is and will be permanently in current interest too. Nevertheless, the problem about control the

robotic wheeled platform is widely researched now, especially in the particulars of controlling the velocities, the accelerations and the paths including their planning. At the same time, optimization the control is not fully researched, especially in considering the different optimum criteria like the used power, the life time, and the jerks. Control of robotic wheeled platforms taking into account the motions' jerks restrictions is one of the not fully developed problem at present, although it really has significant importance for delicate cargo transportation [11], as well as for providing the suitable operation conditions for the sensitive on-board measure devices and systems. It is naturally that considering of the different optimal criteria of the designed control for the robotic wheeled platforms requires the approaches for mathematical modelling and computer simulations relevant with these optimal criteria. Taking into account all these, the subject matter of this research is to develop the theory and methods for computer simulations of the controllability processes of the robotic wheeled platforms, but the goal of this research is in considering the jerks of the wheeled platforms.

To realize this subject matter and goal of this research it is planned to solve the follows tasks:

- developing the generalized approaches for mathematical modelling and computer simulations and theoretical receiving of the properties inherent for the wheeled platforms and suitable for verification the computer simulations results;

- making the calculations to have the quantitative results about the controllability processes for the particular case of the electromechanical four-wheeled platform under the straight motion with the mode of speeding-up from the state of the rest.

To solve the noted tasks it will be used the Lagrange's equations of second kind with the electromechanical analogies, and it will be used final representing of the mathematical models in the form of the system of the first ordered ordinary differential equations with the initial conditions for further numerical solving. The computer simulations will be accomplished by using the Scilab free open source software.

Materials and methods

The general idea of this research is based on using only the computer simulations to have opportunities of imaging the controllability processes for the wheeled platforms without building any physical prototypes. To have the full imagines about the researched controllability processes it is necessary to have the improved generalised approaches providing the relevant mathematical modelling and computer simulations of these processes for the robotic wheeled platforms.

To build the mathematical model of the wheeled platform representing the controllability processes taking into account the different controls' criteria we will assume that the state of the wheeled platform can be defined by the finite dimensional vector \mathbf{x} which will be named as the state vector. It is naturally to imagine that the state vector is depended on the t time

$$\mathbf{x} = \mathbf{x}(t). \quad (1)$$

Besides, we will assume that the state of control influencing on the considered wheeled platform can be defined by the finite dimensional vector \mathbf{u} which will be named further as the vector of the control or just the control. It is naturally to imagine the vector of the control as depended on the time t too:

$$\mathbf{u} = \mathbf{u}(t). \quad (2)$$

Taking into account the assumptions (1) and (2), we can reduce the mathematical modelling of the wheeled platform to building the following mapping:

$$\mathbf{u}(t) \rightarrow \mathbf{x}(t). \quad (3)$$

The mapping (3) shows that some control will lead to some state of the wheeled platform. To research the controllability processes of transition between the different given states of the wheeled platform it is reasonable to define the mapping (3) by means the system of the ordinary differential equations with the initial conditions as it is well-known in the control theory. So, we will assume that the mapping (3) can be represented in the following view:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}; \mathbf{u}), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad (4)$$

where $\mathbf{f}(t, \mathbf{x}; \mathbf{u})$ is the vector defining the velocity of changing the state of the wheeled platform depending on the time t , the current state \mathbf{x} and on the control \mathbf{u} ; t_0 is some given time moment, and \mathbf{x}_0 is the vector defining the state of the researched wheeled platform at the initial time moment $t = t_0$.

The mathematical model (4) of the wheeled platform allows having the full imaginations about changing the state (1) of this platform at the time moments $t \geq t_0$ relevant to some given control (2), and this given control (2) actually characterises the driving power of the wheeled platform. It is possible in general case at least due to numerical solving the Cauchy problem (4), and it is the standard mathematical task, and a lot of computer technologies are existed to solve this task, including the Scilab free open source software suitable for engineering and scientific calculations. At the same time, to design the control systems relevant to different criteria for the robotic wheeled platforms it is necessary to have imagines about time changing of the correspondent parameters defining the controls' criteria. In the case of the controlling the robotic wheeled platforms with the restrictions on the jerk motions formulated for the higher ordered accelerations we will have the following parameters:

$$\mathbf{j} = \Phi \left(\mathbf{x}, \frac{d\mathbf{x}}{dt}, \frac{d^2\mathbf{x}}{dt^2}, \dots \right), \quad (5)$$

where \mathbf{j} is the vector consisted of the different kinds of the jerks used for defining the control's criteria; $\Phi(\dots)$ is some

function relevant to the considered jerks used for defining the control's criteria.

It seems that it is not difficult to calculate the parameters (5) for the given state vector (1), but such calculating has in fact some principal difficulties due to involving the derivatives. Really, computing the derivatives required to estimate the criteria's parameters (5) is the principal problem because of it we will have no the exact analytical solutions for the state vector (1), but we will have only the approximate solution for this vector (1), and besides, this approximate solution will be represented in the discrete form of the values corresponded to the time moments relevant to the integrating steps. It seems, that it is possible to use the finite differences technique to find the derivatives of the discrete defined functions, but the discrete definitions of the functions will be predefined by the integrating procedure, so to have opportunities of using the finite differences technique the we must have the relevant integrating steps to have the suitable representation of the state vector (1) as the result of numerical solving the Cauchy problem (4). At the same time, it is well-known, differentiating of the approximate functions is the incorrect problem in the Hadamard's sense, because of the small changing of the function will lead to the big and even infinite changing of their derivatives, so using the finite differences technique to find the derivatives of the state vector (1) by using its approximate solution in any case will lead to the incorrect results, even if we will have the reliable grid of the data. We had experience of viewing such incorrectness in the previous researches [10, 11], so ours previous researches has no results about the higher derivatives of the state vector (1). At the same time, the higher derivatives of the state vector (1) are significantly required to estimate the some kinds of the quality of the control, including the jerks for example which are defined as the time derivative of the acceleration. Thus, it is significantly required to have the reliable general approach for evaluating the derivatives of the state vector (1) involved in the definition (5) of the parameters defining the quality of the control. The main principle of the reliable general approach for evaluating the derivatives of the state vector (1) involved in the definition (5) is in excluding the differentiations of the approximate defined state vector (1) by using the differential equations (4):

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}, \quad \frac{d^2\mathbf{x}}{dt^2} = \frac{\partial \mathbf{f}}{\partial t} + \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \mathbf{f} + \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{dt}, \dots, \quad (6)$$

where $\mathbf{f} = \mathbf{f}(t, \mathbf{x}; \mathbf{u})$.

The relations (6) based on the mathematical model (4) allow having the correct evaluations for the derivatives of the state vector (1) represented by the approximate solution of the Cauchy problem (4), because of excluding the differentiation of the approximately represented functions with the incorrectness in the Hadamard's sense. Really, the partial derivatives $\partial \mathbf{f} / \partial t$, $\partial \mathbf{f} / \partial \mathbf{x}$ and $\partial \mathbf{f} / \partial \mathbf{u}$ can be find correctly due to the exactly defined mathematical model (4), but the derivative $d\mathbf{u} / dt$ can be find correctly because of existing the exact data about the control. It is necessary to note, that evaluating the third

and higher derivatives using the scheme (6) will lead to the more cumbersome relations, but all these relations can be realized correctly.

Using the mathematical modelling (4) especially effective is combining with the computer simulations providing automatically the required numerical solving of the Cauchy problem (4) and the follows calculations using the relations (6) for evaluating the researched parameters (5). It is naturally that using such computer simulations requires the reliable substantiations of correctness in using the relevant especially developed software or the proposed computer models developed by the existed tools of the used software similar to the Scilab, Matlab, and others. Such substantiations are often reduced to considering the especially formulated tasks with the known solutions and to comparing the results of the computer simulations with these known solutions. To substantiate the correctness of the computer simulations based on using the mathematical model (4) it is proposed to use the test task corresponded to the following case:

$$t_0 = 0, \mathbf{x}_0 = \mathbf{0}, \mathbf{u}(t) = \mathbf{u}_c, \quad (7)$$

where $\mathbf{0}$ is the zero vector with the relevant dimension; \mathbf{u}_c is some given constant vector.

The well-known fundamental property inherent for the wheeled platforms is in existing of the maximal velocity corresponded to the equilibrium between the driving and resistance powers, so taking into account the relations (7), we will have the follows:

$$\lim_{\substack{t \rightarrow \infty, \\ u = u_c}} \frac{d\mathbf{x}}{dt} = \mathbf{v}_c \Rightarrow \lim_{\substack{t \rightarrow \infty, \\ u = u_c}} \mathbf{x}(t) = \mathbf{v}_c t, \quad (8)$$

where \mathbf{v}_c is some constant vector corresponded to the \mathbf{u}_c is some given constant vector defining the control.

The mathematical model (4) and the properties (8) allow to formulate the equations for evaluating the \mathbf{v}_c vector corresponded to the given constant vector \mathbf{u}_c defining the control:

$$\lim_{t \rightarrow \infty} \mathbf{f}(t, \mathbf{v}_c t, \mathbf{u}_c) = \mathbf{v}_c. \quad (9)$$

Thus, the relation (9) is the consequence of the mathematical model (4), and the results of the computer modelling the robotic wheeled platforms and their controllability processes must be in fully agreement with

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_k} - \frac{\partial \mathcal{L}}{\partial q_k} = -\frac{\partial \mathcal{R}}{\partial \dot{q}_k} - \frac{\partial \mathcal{H}}{\partial \dot{q}_k} + Q_k(t, q_1, q_2, \dots, q_n, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n; \mathbf{u}), \quad k = 1, 2, \dots, n, \quad (11)$$

where \mathcal{L} is the Lagrange's function defined as the difference between the kinetic and potential energies; \mathcal{R} is the Raleigh's function defining the damping; \mathcal{H} is the function defining the gyroscopic effects; Q_k is the generalized force relevant to the generalized coordinate q_k .

In the Lagrange's equations of second kind represented in the form (11) it is assumed that the generalized coordinates Q_k represent the driving power of

the relation (9). The relation (9) represents the most general property of the wheeled platforms and it can be sufficient for substantiating the correctness of the computer simulations. It is necessary to note, that the considered particular case (7) corresponds to the case of the transition function of the robotic wheeled platform imagined as the automation object, and it is in agreement with the purposes of this research in developing the approaches for computer simulations of the controllability processes.

Representing the mathematical models of the robotic wheeled platforms in the generalized view (4) is in agreement with the common approaches of the mathematical modelling of the discrete systems, but it is necessary to have the approaches to build the mathematical model (4) representing the particular robotic wheeled platform. It is really difficult to give the really generalized approaches for building the mathematical model (4) for the particular robotic wheeled platform, but it is naturally to use the well-known approaches of analytical mechanics and electromechanical analogies, because of the robotic wheeled platforms are usually the electromechanical systems. We will assume further the robotic wheeled platforms as the holonomic systems, because of such schematisation is suitable for considering some operational modes including the straight motion on their rigid road [10, 11]. This assumption allows using the principal idea of the analytical mechanics about the freedoms degrees in defining the state of the wheeled platforms by means the independent parameters:

$$q_k = q_k(t), k = 1, 2, \dots, n, \quad (10)$$

where q_k is the generalised coordinate; n is the number of the freedom degrees for the considered wheeled platforms.

It should be noted that the generalized coordinates (10) can represent not only the mechanical values like the linear and curvilinear displacements as well as the angles of the rotations, but also the electrical values like the electric charges or the voltages in depending on the used electromechanical analogies for modelling the electromechanical systems. In any case, it is well-known in the analytical mechanics that the most general representation of the differential equations representing the holonomic electromechanical systems is based on using the electromechanical analogies and the Lagrange's equations of second kind:

the considered wheeled platform. The Lagrange's equations of second kind (11) represent the system of the second ordered ordinary differential equations, so these equations must be considered with the relevant initial conditions:

$$q_k(t_0) = q_k^{(0)}, \dot{q}_k(t_0) = \dot{q}_k^{(0)}, \quad k = 1, 2, \dots, n, \quad (12)$$

where $q_k^{(0)}$ and $\dot{q}_k^{(0)}$ are the given initial values of the generalized coordinates and the generalized velocities.

The generalized coordinates (10) allow finding different parameters representing the different characteristics of the controlled wheeled platforms. To consider the controllability processes for the robotic wheeled platforms under the restrictions on the motions' jerks it is required considering the quantitative evaluations of these jerks which will be considered further as the time derivative of the acceleration of the mass center of the wheeled platform. In this research, we will assume that the mass center has the straight motion, so for the wheeled platform considering as the holonomic system with the generalized coordinates (10) we will have the following:

$$s = s(q_1, q_2, \dots, q_n), \quad (13)$$

where s is the natural coordinate of the mass center of the wheeled platform.

Due to assumption about the straight motion of the mass center, and due to the assumed quantitative evaluation of the jerks as the time derivative of the acceleration, the relation (13) allows having the follows representation of the jerk for the considered robotic wheeled platform:

$$j(t) = \frac{d^3 s}{dt^3} \Rightarrow j(t) = \sum_{k=1}^n \left(\frac{\partial^3 s}{\partial q_k^3} \left(\frac{dq_k}{dt} \right)^3 + 3 \frac{\partial^2 s}{\partial q_k^2} \frac{dq_k}{dt} \frac{d^2 q_k}{dt^2} + \frac{\partial s}{\partial q_k} \frac{d^3 q_k}{dt^3} \right), \quad (14)$$

where $j(t)$ is the quantitative evaluation of the jerk.

Relation (14) shows that the jerks of the translational motions of the wheeled platforms are depended on the generalized velocities, generalized accelerations and the time derivatives of the generalized accelerations, as well as on the building of the wheeled platform.

It is necessary to understand that the Lagrange's equations of second kind (11) with the initial conditions (12) have not direct form (4), but it is possible to represent the differential equations (11) and the initial conditions (12) in the form (4) by means introducing the relevant state vector (1) in the following view:

$$\begin{aligned} \mathbf{x}^T &= (x_1 \quad x_2 \quad \dots \quad x_n \quad x_{n+1} \quad x_{n+2} \quad \dots \quad x_{2n}), \\ x_1 &= q_1, x_2 = q_2, \dots, x_n = q_n, \\ x_{n+1} &= \dot{q}_1, x_{n+2} = \dot{q}_2, \dots, x_{2n} = \dot{q}_n. \end{aligned} \quad (15)$$

Although, it is possible principally to consider the mathematical model representing the wheeled platform in the view (12), (13) of the system of the second ordered differential equations, but the most of numerical method for solving the Cauchy problems are foresaw for the systems of the first ordered differential equations (4), so reducing the problem (12), (13) to the equivalent generalized form (4) is principally required for computer simulations, and the Lagrange's equations of second kind are just the approach for building of the mathematical models of the robotic wheeled platforms. Thus, we have the generalised approaches to build the mathematical models of the controllability processes of the robotic wheeled platforms based on the Lagrange's equations of second kind (11) and the electromechanical analogies well-known in the analytical mechanics. Besides, reducing the relevant Lagrange's equations of second kind (11) to the equivalent system (4) of the first ordered differential equations as well as using the differential equations to evaluate the derivatives (6) of the state vector all allow providing the computer simulations of the controllability processes for the robotic wheeled platforms taking into account the restrictions of the motion's jerks, and due to the relation (9) we have the reliable criteria for substantiating the correctness of the results of such computer simulations.

Study results and their discussion

To illustrate the possibilities of the proposed generalized approaches and to have the quantitative estimations about the controllability processes in the robotic wheeled platforms we will consider the particular case of the four-wheeled electromechanical robotic platform, because this case covers the wide class of the robotic wheeled platforms typical for the different applications. Schematisation of the considered four-wheeled electromechanical robotic platform is presented on the fig. 1. It is assumed, that the considered wheeled platform has the straight translational motion of the housing (pos. 1 on fig. 1a) defining by the s natural coordinate, and this motion is due to rotations of two driving wheels (pos. 2 on fig. 1a) defining by the rotation angle φ and occurred by the drive direct current electric motors joined with this wheels (pos. 3 on fig. 1a); the housing (pos. 1 on fig. 1a) of the wheeled platform is supported on the road by means two driven wheels (pos. 4 on the fig. 1a). The viscous damping on the housing will be represented be the force F_v depending on the velocity of the translational motion, and the rolling friction will be represented by the constant rolling friction couples M_{π} on the wheels. The equivalent scheme of the drive electric motors (fig. 1b) is reduced to the electrical inductance, the electrical resistance, to the supplied voltage U and to the induced voltage depending on the rotor angle velocity and counteracting to the supplied voltage. The state of the electrical parts of the driving direct current electric motors is defined by the electric charge e (fig. 1b). To have opportunities to show the correctness of the results of the computer simulations we will consider the case corresponded to the relations (7)–(9), so we will consider the particular case of the initial state corresponded to the state of the rest and the case of the constant supplied voltage:

$$U(t) = U_C, \quad (16)$$

where U_C is some given constant.

To make the reliable computer simulations of the controllability processes for the considered particular

robotic wheeled platform (fig. 1) it is firstly necessary to build the relevant mathematical model, and to do this we will use the Lagrange's equations of second kind (11) with the electromechanical analogies as it was foresaw in the proposed generalised approaches. We will have the following evident relation between the translational

coordinate of the housing and the rotation angles of all wheels (fig. 1a):

$$s = r\varphi, \quad (17)$$

where r is the radius of the wheels (fig. 1a).

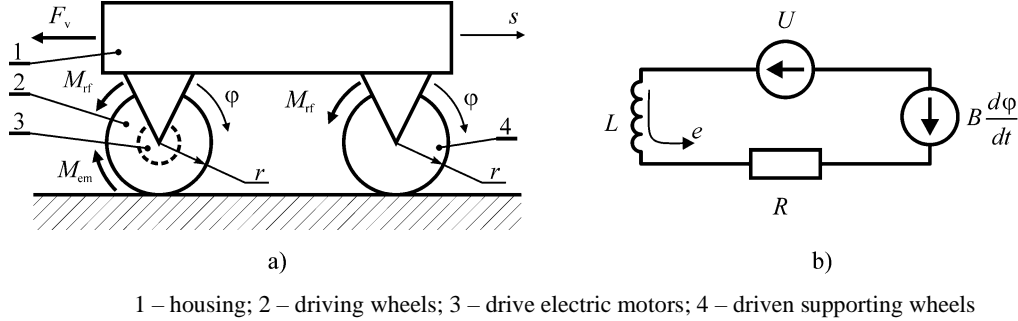


Fig. 1. Schematisation of the mechanics (a) and the electric motors (b) of the considered wheeled platform

Due to the relation (17) we have the holonomic system with two freedoms' degrees ($n = 2$), and we will use the follows generalized coordinates:

$$q_1 = \varphi, q_2 = e. \quad (18)$$

So, taking into account the generalised coordinates (18) and the relation (17) the Lagrange's function \mathcal{L} , the

$$\mathcal{L} = \frac{1}{2} J \dot{q}_1^2 + \frac{1}{2} 2L \dot{q}_2^2, \mathcal{R} = 4M_{rf} \dot{q}_1 + \frac{1}{2} \beta r^2 \dot{q}_1^2 + \frac{1}{2} 2R \dot{q}_2^2, \mathcal{H} = 0, Q_1 = 2M_{em}, Q_2 = 2(U_c - B \dot{q}_1). \quad (19)$$

where J is the inertia moment of the wheeled platform relatively the rotation of the wheels; β is the parameter defining the linear viscous damping proportional to the velocity of the housing (pos. 1 on fig. 1a); L is the inductance, R is the resistance of the equivalent electric circuit and B is the electromechanical parameter of the driving direct current electric motor.

We will assume that the driving couple from the direct current electric motor is proportional to the electric current in the drive's electric circuit

$$M_{em} = B \dot{q}_2. \quad (20)$$

The Lagrange's equations of second kinds (11) and the relations (19), (20) as well as the assumptions about the state of the rest at the initial time moment will lead to the follows differential equations:

Raleigh's function \mathcal{R} , the function \mathcal{H} defining the gyroscopic effects, and the generalized forces Q_k for the considered electromechanical wheeled platform (fig. 1) will have the following views:

$$\begin{aligned} J \ddot{q}_1 &= -\beta r^2 \dot{q}_1 + 2B \dot{q}_2 - 4M_{rf}, \\ 2L \ddot{q}_2 &= -2B \dot{q}_1 - 2R \dot{q}_2 + 2U_c \end{aligned} \quad (21)$$

$$q_1(0) = 0, \dot{q}_1(0) = 0, q_2(0) = 0, \dot{q}_2(0) = 0. \quad (22)$$

We have the mathematical model (21), (22) representing the considered robotic wheeled platform (fig. 1), but to make computer simulations of the controllability processes we must represent this mathematical model (21), (22) in the form (4) of the system of the first ordered differential equations. To do this, we will use the new variables (15) corresponded to the case of the $n = 2$. Besides, we will use the control vector (2) as the one dimensions vector with one element represented by the constant voltage (16). Taking into account all these circumstances, we will have the follows vectors representing the problem (21), (22) in the suitable form (4):

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \mathbf{u} = (U_c), \mathbf{f}(t, \mathbf{x}, \mathbf{u}) = \begin{pmatrix} x_3 \\ x_4 \\ -\frac{\beta r^2}{J} x_3 + \frac{2B}{J} x_4 - \frac{4M_{rf}}{J} \\ -\frac{B}{L} x_3 - \frac{R}{L} x_4 + \frac{U_c}{L} \end{pmatrix}, t_0 = 0, \mathbf{x}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (23)$$

The generalized relations (13) and (14) representing the natural coordinate and the jerk, as well as the relations for the velocity and accelerations of the straight motion and for the electric current in the drive electric motors of

the considered wheeled platform taking into account the relations (17), (18) (15) will have the following view:

$$s = r x_1, v = r x_3, a = r \frac{dx_3}{dt}, j = r \frac{d^2 x_3}{dt^2}, I = x_4, \quad (24)$$

where s is the natural coordinate, v is the velocity, a is the acceleration, j is the jerk of the straight motion of the mass center and I is the electric current in the drive electric motors of the considered wheeled platform.

$$\frac{dx_3}{dt} = -\frac{\beta r^2}{J} x_3 + \frac{2B}{J} x_4 - \frac{4M_{rf}}{J}, \quad \frac{d^2 x_3}{dt^2} = -\frac{\beta r^2}{J} \left(-\frac{\beta r^2}{J} x_3 + \frac{2B}{J} x_4 - \frac{4M_{rf}}{J} \right) + \frac{2B}{J} \left(-\frac{B}{L} x_3 - \frac{R}{L} x_4 + \frac{U_c}{L} \right). \quad (25)$$

Thus, we have the mathematical model of the researched controllability processes for the considered robotic wheeled platform (fig. 1) in the suitable view (4), (23) for the computer simulations, and we can define wished additional parameters characterizing the quality of the control by means of the correct formulas (24), (25).

To have substantiations about reliability the results of computer simulation we will use the relation (9). In the considered example (fig. 1) the steady state (8) corresponded to the equilibrium between the driving and the damping powers for the considered wheeled platform (fig. 1) will be corresponded to the motion of the wheeled platform with the constant velocity and to the constant electric current in the electric motors circuits:

$$\mathbf{v}_c^T = (v_c \quad I_c \quad 0 \quad 0), \quad (26)$$

where v_c is the maximum velocity possible for the considered wheeled platform and I_c is the relevant constant current in the drive electric motors' circuits under the U_c supplied voltage.

The relation (9) for the vector (26), the relations (23) defining the problem (4) will lead to the following:

$$-\beta r v_c + 2BI_c = 4M_{rf}, \quad \frac{B}{r} v_c + RI_c = U_c. \quad (27)$$

The general relations (6) allow excluding the differentiation of the state vector from the formulas (24) in the considered particular cases will have the following view:

Relations (27) can be considered as the system of two linear algebraic equations for finding the established constant velocity v_c of the wheeled platform and the relevant electric current I_c in the drive electric motors' circuit:

$$v_c = r \frac{2BU_c - 4M_{rf}R}{\beta r^2 R + 2B^2}, \quad I_c = \frac{\beta r^2 U_c + 4BM_{rf}}{\beta r^2 R + 2B^2}. \quad (28)$$

Exactly the values (28) will be used for comparing with the results of computer simulations to show correctness of these results.

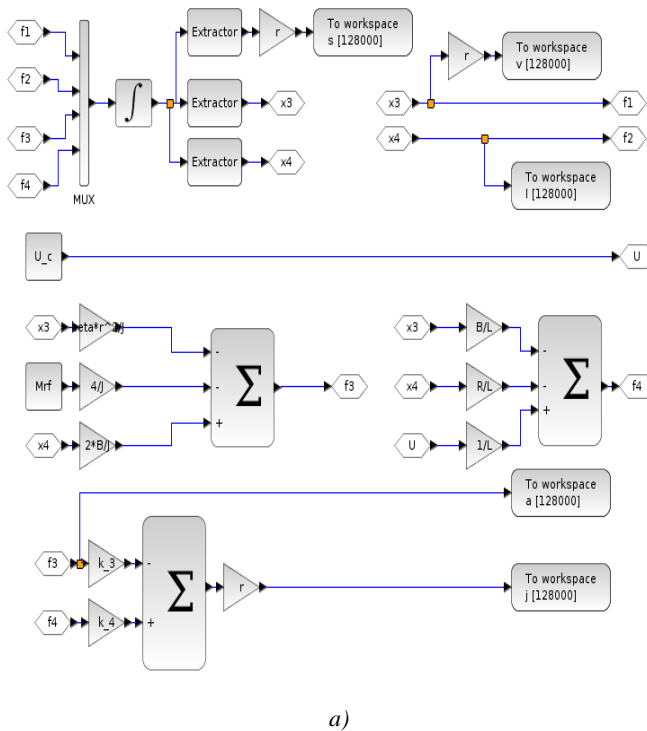
We will consider further the computer simulations of the controllability processes for the considered wheeled platform (fig.1) with the follows numerical parameters:

$$J = 80 \text{ kg} \cdot \text{m}^2, r = 0, 2 \text{ m}, \beta = 75 \text{ kg/s},$$

$$M_{rf} = 50 \text{ N} \cdot \text{m}, L = 2, 6 \text{ mH}, R = 1, 18 \Omega, \quad (29)$$

$$B = 4 \text{ N} \cdot \text{m/A}$$

To direct accomplish such computer simulations of the controllability processes for the considered robotic wheeled platforms (fig. 1) based on the mathematical model (4), (23), on the relations (24), (25) and (28), and on the input data (29) we will use the Scilab free open source software as shown on the fig. 2. The results of the computer simulations are presented at fig. 3 and fig. 4.



```

1 J=80; r=0.2; beta=75; Mrf=50; L=2.6E-3; R=1.18; B=4; U_c=60;
2 k_3=beta*r^2/J; k_4=2*B/J
3 loadXcosLibs();
4 importXcosDiagram("/home/yurii/MyScilab/SP2022_F/SP2022_F4.zcos");
5 typeofs(scs_m); scs_m.props.context;
6 scicos_simulate(scs_m);
7 Vc=r*(2*B*U_c-4*Mrf*R)/(beta*r^2*R+2*B^2);
8 Ic=(beta*r^2*U_c+4*B*Mrf)/(beta*r^2*R+2*B^2);
9 subplot(3,2,1); plot(s.time,s.values,'black');
10 subplot(3,2,2); plot(v.time,v.values,'black');
11 subplot(3,2,2); plot([0 20],[Vc Vc],'black');
12 subplot(3,2,3); plot(a.time,a.values,'black');
13 subplot(3,2,4); plot(j.time,j.values,'black');
14 subplot(3,2,5); plot(I.time,I.values,'black');
15 subplot(3,2,5); plot([0 20],[Ic Ic],'black');
16 U_c=50; scicos_simulate(scs_m);
17 Vc=r*(2*B*U_c-4*Mrf*R)/(beta*r^2*R+2*B^2);
18 Ic=(beta*r^2*U_c+4*B*Mrf)/(beta*r^2*R+2*B^2);
19 subplot(3,2,1); plot(s.time,s.values,'red');
20 subplot(3,2,2); plot(v.time,v.values,'red');
21 subplot(3,2,2); plot([0 20],[Vc Vc],'red');
22 subplot(3,2,3); plot(a.time,a.values,'red');
23 subplot(3,2,4); plot(j.time,j.values,'red');
24 subplot(3,2,5); plot(I.time,I.values,'red');
25 subplot(3,2,5); plot([0 20],[Ic Ic],'red');
26 U_c=40; scicos_simulate(scs_m);
27 Vc=r*(2*B*U_c-4*Mrf*R)/(beta*r^2*R+2*B^2);
28 Ic=(beta*r^2*U_c+4*B*Mrf)/(beta*r^2*R+2*B^2);
29 subplot(3,2,1); plot(s.time,s.values,'blue');
30 subplot(3,2,2); plot(v.time,v.values,'blue');
31 subplot(3,2,2); plot([0 20],[Vc Vc],'blue');
32 subplot(3,2,3); plot(a.time,a.values,'blue');
33 subplot(3,2,4); plot(j.time,j.values,'blue');
34 subplot(3,2,5); plot(I.time,I.values,'blue');
35 subplot(3,2,5); plot([0 20],[Ic Ic],'blue');

```

Fig. 2. Graphical representing of the computer model (a) and the script (b) for automation on the basis of the program, control of the researches in the Scilab free open source software

The results of the computer simulations (fig. 3) are in full agreement with the fundamental property of the wheeled platforms generally represented by the relations (7)–(9) which had the particular view (26)–(28) for the considered wheeled platform (fig. 1). This full agreement shows correctness of the result of the computer simulations using the computer models and script (fig. 2) especially developed using the Scilab tools. The results for the jerks (fig. 4) show that the maximum value of the jerk is at the initial time moment, and this is in full agreement with the physical sense of the considered problem about speeding-up the wheeled platform from the state of the rest. Considering the results for the different values of the voltage (16) supplied on the drive electric motors (fig. 3 and fig. 4) allows having imagination about the controllability of the considered wheeled platform and it allows having the imagination about this wheeled platform as the automation object. We can see (fig. 4) the

significant values of the jerks and these results are due to the impact supplying of the voltage on the drive electric motors of the wheeled platform. It is possible to decrease the jerk by decreasing the value of the supplied voltage, but such smaller values of the jerks will be significant yet, so to significantly decrease the jerk it is necessary to provide the smooth time dependent of the voltage supplied on the drive electric motors. The significant values of the electric current in the winding on the drive electric motors at the initial time moment (fig. 3d) are due to the impact supplying of the voltage on the drive electric motors too. All these results (fig. 3 and fig. 4) of the computer simulations lead to the conclusion that optimisation of control by means choosing the suitable smooth voltage supplied on the drive electric motors will allow increasing the quality of operation of the considered wheeled platform.

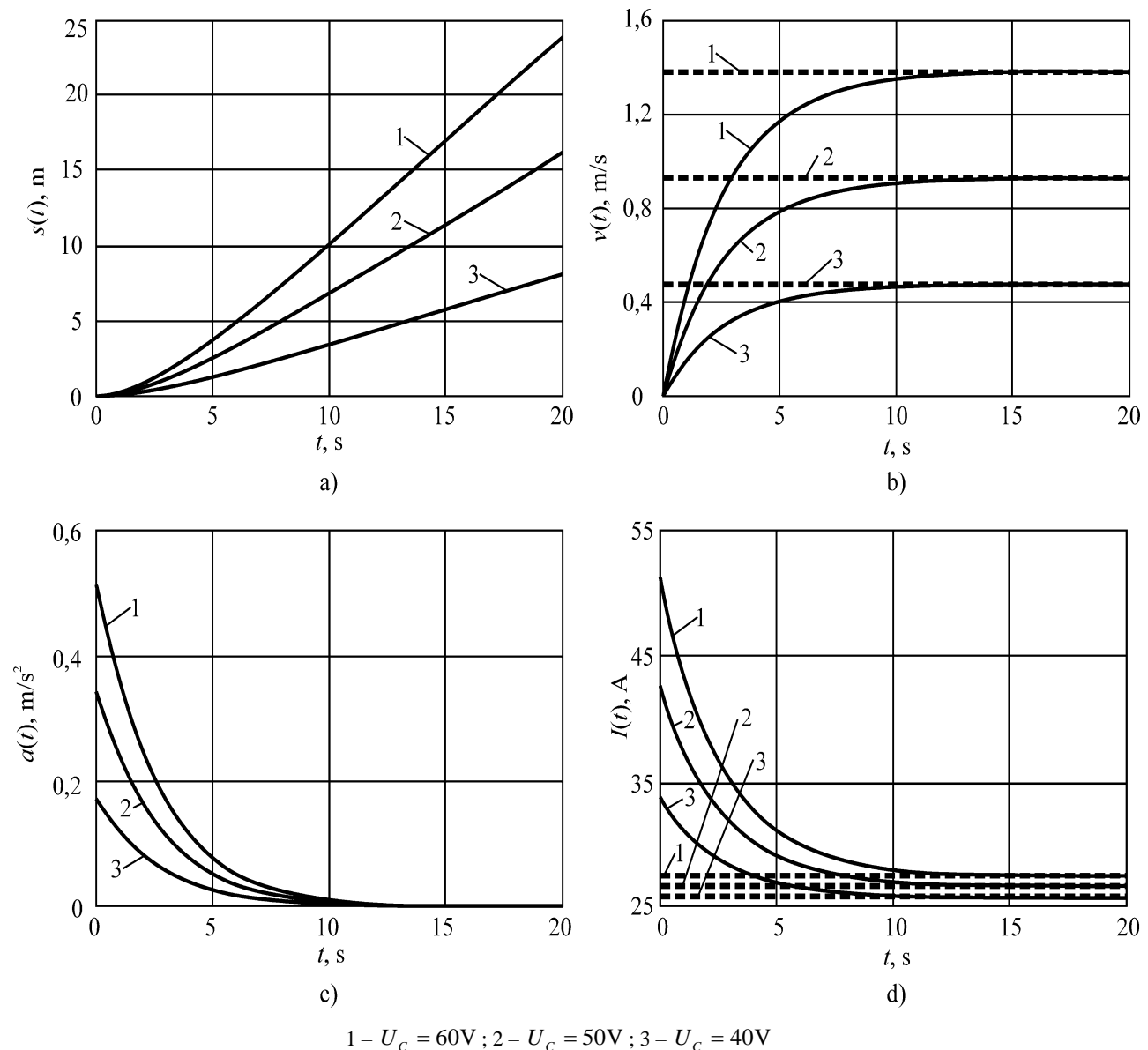


Fig. 3. Results of computer simulations (solid) of the translational coordinate (a), the velocity (b), the acceleration (c) and for the electric current (d) in the drive electric motors for the robotic wheeled platform, as well as the theoretically predicted results (dash) for the steady state

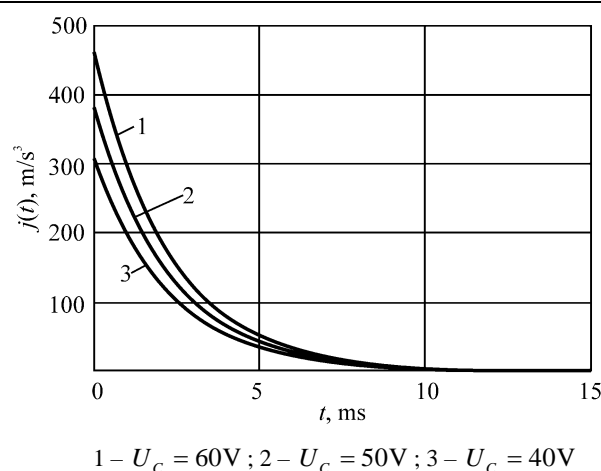


Fig. 4. Results of computer simulations of the jerks for the robotic wheeled platform

Conclusion and perspectives of further development

Successful accomplishing of this research allows formulating the conclusions important for developing the theory and methods for computer simulations of the controllability processes of the robotic wheeled platforms.

It is developed the generalized approaches for computer simulations of the controllability processes of the robotic wheeled platforms, and these approaches have the advantages comparing with the existed due to possibilities of correct estimating not only the coordinates, the velocity, the acceleration, but as well as the jerk motions which are required for designing the control optimal in the different senses. These approaches are based on the mathematical modeling of the robotic wheeled platforms by using the Lagrange's equations of second kind with the electromechanical analogies, so the wide kinds of the robotic electromechanical wheeled platforms can be considered. The computer simulations of the controllability processes for the robotic wheeled platforms are reduced to numerical solving of the differential equations and the initial conditions of the relevant mathematical models represented in the suitable form of the system of first ordered ordinary differential equations with the initial conditions. On the example of the Scilab free open source software it is shown that the universal computer systems developed for the scientific and engineering calculations can be effectively used for computer simulations of the controllability processes for the robotic wheeled platforms.

It is proposed to use the differential equations of the mathematical model to compute the second and higher derivatives of the state vector necessary for evaluating the motions' jerks of the wheeled platforms using the numerical approximate solutions. This proposed approach

allows excluding the incorrectness in the Hadamard's sense due to necessities of differentiating of the approximate functions. To have the reliable substantiations of correctness of the results of computer simulations of the controllability processes for the robotic wheeled platform it is proposed to use the property inherent for the wheeled platforms in existing of the maximal steady velocity relevant to the equilibrium between the driving and damping powers.

In the considered particular case of the four-wheeled electromechanical robotic platform typical for the different applications it is shown, that the results of the computer simulations are in full agreement with the fundamental property inherent for the wheeled platforms, so it shows correctness of the result of the computer simulations using the computer models and the script especially developed using the Scilab tools. It is shown that the maximum value of the jerk is at the initial time moment, and this is in full agreement with the physical sense of the considered problem about speeding-up the wheeled platform from the state of the rest. It is shown also, that considering the results for the different values of the voltage supplied on the drive electric motors allows having the imaginations about the controllability of the wheeled platforms and it allows having the imagination about this wheeled platform as the automation object. The significant values of the jerks are due to the impact supplying of the voltage on the drive electric motors of the wheeled platform. It is shown that optimization of control by means choosing the suitable smooth voltage supplied on the drive electric motors will allow increasing the quality of operation of the considered wheeled platform. For the further it is recommended to research the controllability processes for other parameters like the supplied power and others.

References

1. Houtman, W., Lopez Martinez, C.A., Wang, S., Ketels, A., Bruyninckx, H.P.J., van de Molengraft, M.J.G. (2021), "Dynamic control of steerable wheeled mobile platforms applied to an eight-wheeled RoboCup Middle Size League soccer robot", *Mechatronics*, Vol. 80, 102693. DOI: <https://doi.org/10.1016/j.mechatronics.2021.102693>
2. Peng, G., Lu, Z., Tan, Z., He, D., Li, X. (2021), "A novel algorithm based on nonlinear optimization for parameters calibration of wheeled robot mobile chasses", *Applied Mathematical Modelling*, Vol. 95, P. 396–408. DOI: <https://doi.org/10.1016/j.apm.2021.02.012>

3. Moreno-Caireta, I., Celaya, E., Ros, L. (2021), "Model Predictive Control for a Mecanum-wheeled Robot Navigating among Obstacles", *IFAC-PapersOnLine*, Vol. 54 (6), P. 119–125. DOI: <https://doi.org/10.1016/j.ifacol.2021.08.533>
4. Sun, Z., Hu, S., He, D., Zhu, W., Xie, H., Zheng, J. (2021), "Trajectory-tracking control of Mecanum-wheeled omnidirectional mobile robots using adaptive integral terminal sliding mode", *Computers & Electrical Engineering*, Vol. 96 (A), 107500. DOI: <https://doi.org/10.1016/j.compeleceng.2021.107500>
5. Mondal, S., Ray, R., Reddy N. S., Nandy, S. (2022), "Intelligent controller for nonholonomic wheeled mobile robot: A fuzzy path following combination", *Mathematics and Computers in Simulation*, Vol. 193, P. 533–555. DOI: <https://doi.org/10.1016/j.matcom.2021.10.028>
6. Jiang, L., Wang, S., Xie, Y., Quan Xie, S., Zheng, S., Meng, J. (2022), "Fractional robust finite time control of four-wheel-steering mobile robots subject to serious time-varying perturbations", *Mechanism and Machine Theory*, Vol. 169, 104634. DOI: <https://doi.org/10.1016/j.mechmachtheory.2021.104634>
7. Feng Li, Z., Tao Li, J., Fan Li, X., Jian Yang, Y., Xiao, J., Wen Xu, B. (2020), "Intelligent Tracking Obstacle Avoidance Wheel Robot Based on Arduino", *Procedia Computer Science*, Vol. 166, P. 274–278. DOI: <https://doi.org/10.1016/j.procs.2020.02.100>
8. Alipour, K., Robat, A.B., Tarvizadeh, B. (2019), "Dynamics modeling and sliding mode control of tractor-trailer wheeled mobile robots subject to wheels slip", *Mechanism and Machine Theory*, Vol. 138, P. 16–37. DOI: <https://doi.org/10.1016/j.mechmachtheory.2019.03.038>
9. Bayar, G., Ozturk, S. (2020), "Investigation of The Effects of Contact Forces Acting on Rollers Of a Mecanum Wheeled Robot", *Mechatronics*, Vol. 72, 102467. DOI: <https://doi.org/10.1016/j.mechatronics.2020.102467>
10. Mamalis, A.G., Nevliudov, I., Romashov, Yu. (2021), "An approach for numerical simulating and processing of measured electrical signals from board sensors installed on wheeled electro-mechanical platforms", *Journal of Instrumentation*, Vol. 16 (10), P10006. DOI: <https://doi.org/10.1088/1748-0221/16/10/P10006>
11. Alyokhina, S., Nevliudov, I., Romashov, Y. (2021), "Safe Transportation of Nuclear Fuel Assemblies by Means of Wheeled Robotic Platforms", *Nuclear and Radiation Safety*, Vol. 3 (91), P. 43–50. DOI: [https://doi.org/10.32918/nrs.2021.3\(91\).05](https://doi.org/10.32918/nrs.2021.3(91).05)

Received 31.03.2022

Відомості про авторів / Сведения об авторах / About the Authors

Альошина Світлана Вікторівна – доктор технічних наук, старший науковий співробітник, Інститут проблем машинобудування ім. А.М. Підгорного НАН України, в.о. завідувача відділу оптимізації процесів і конструкцій турбомашин; Харківський національний університет радіоелектроніки, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації та мехатроніки, м. Харків, Україна; e-mail: alyokhina@ipmach.kharkov.ua; ORCID ID: <https://orcid.org/0000-0002-2967-0150>.

Алехина Светлана Викторовна – доктор технических наук, старший научный сотрудник, Институт проблем машиностроения им. А.М. Подгорного НАН Украины, и.о. заведующего отделом оптимизации процессов и конструкций турбомашин; Харьковский национальный университет радиоэлектроники, профессор кафедры компьютерно-интегрированных технологий, автоматизации и мехатроники, г. Харьков, Украина.

Alyokhina Svitlana – Doctor of Engineering Sciences, Senior Researcher, A. Pidhomy Institute of Mechanical Engineering Problems of the National Academy of Sciences of Ukraine, Acting Head of Department of Optimization of Processes and Designs of Turbomachines; Kharkiv National University of Radioelectronics, Professor of the Department of Computer-Integrated Technologies, Automation and Mechatronics, Kharkiv, Ukraine.

Невлиудов Ігор Шакирович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри комп'ютерно-інтегрованих технологій, автоматизації та мехатроніки, м. Харків, Україна; e-mail: igor.nevliudov@nure.ua; ORCID ID: <https://orcid.org/0000-0002-9837-2309>.

Невлиудов Игорь Шакирович – доктор технических наук, профессор, Харьковский национальный университет радиоэлектроники, заведующий кафедрой компьютерно-интегрированных технологий, автоматизации и мехатроники, г. Харьков, Украина.

Nevliudov Igor – Dr. Sc. (Engineering), Professor, Kharkiv National University of Radioelectronics, Head the Department of Computer-Integrated Technologies, Automation and Mechatronics, Kharkiv, Ukraine.

Ромашов Юрій Володимирович – доктор технічних наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації та мехатроніки; Харківський Національний Університет імені В.Н. Каразіна, професор кафедри прикладної математики, м. Харків, Україна; e-mail: yurii.romashov@nure.ua; ORCID ID: <https://orcid.org/0000-0001-8376-3510>.

Ромашов Юрий Владимирович – доктор технических наук, доцент, Харьковский национальный университет радиоэлектроники, профессор кафедры компьютерно-интегрированных технологий, автоматизации и мехатроники; Харьковский национальный университет имени В.Н. Каразина, профессор кафедры прикладной математики, г. Харьков, Украина.

Romashov Yuri – Doctor of Engineering Sciences, Docent, Kharkiv National University of Radio Electronics, Professor of the Department of Computer-Integrated Technologies, Automation and Mechatronics; V.N. Karazin Kharkiv National University, Professor of the Applied Mathematics Department, Kharkiv, Ukraine.

КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ ПРОЦЕСІВ КЕРОВАНOSTІ ДЛЯ РОБОТИЗОВАНИХ КОЛІСНИХ ПЛАТФОРМ З ВРАХУВАННЯМ ОБМЕЖЕНЬ РИВКІВ РУХІВ

Комп'ютерне моделювання розглядається як необхідний інструмент для розробки відповідних автономних систем управління, оптимальних у різних сенсах, і, особливо, щодо забезпечення обмежень ривків рухів для роботизованих колісних платформ. **Предметом** дослідження є розробка теорії та методики комп'ютерного моделювання процесів

керованості роботизованих колісних платформ. **Метою** даного дослідження є врахування ривків колісних платформ, які визначаються похідними за часом прискорення центру мас колісної платформи, так що ця похідна розглядається як кількісна оцінка ривків. Обговорюється некоректність у розумінні Адамара визначення ривків шляхом прямого диференціювання фазових координат у разі комп'ютерного моделювання за допомогою обчислювальних методів. **Завдання** даного дослідження полягають у розробці узагальнених підходів щодо математичного та комп'ютерного моделювання та в теоретичному отриманні властивостей колісних платформ, що придатні для перевірки результатів комп'ютерного моделювання, а також у проведенні розрахунків для отримання кількісних результатів щодо процесів керованості для конкретного випадку електромеханічної чотириколісної платформи при прямолінійному русі в режимі розганяння зі стану спокою. **Методи** цього дослідження базуються на рівняннях Лагранжа другого роду, а також на електромеханічних аналогіях та на остаточному представленні математичних моделей у вигляді системи звичайних диференціальних рівнянь першого порядку з початковими умовами для подальшого обчислювального розв'язування. Комп'ютерне моделювання виконується за допомогою вільного програмного забезпечення Scilab з відкритим кодом. **Результати** цього дослідження полягають у запропонованому коректному способі обчислення ривків через фазові координати без їхнього диференціювання, що дозволяє виключити некоректність у розумінні Адамара, а також представити процеси керованості для електромеханічної колісної платформи, включаючи результати для швидкості, прискорення та ривків, які необхідні для ілюстрації процесів керованості роботизованих колісних платформ. Шляхом порівняння з теоретично встановленими властивостями колісних платформ показано правильність результатів комп'ютерного моделювання. **Висновки** щодо цього дослідження полягають у тому, що розроблені підходи щодо комп'ютерного моделювання процесів керованості роботизованих колісних платформ дозволяють врахувати вплив керування на різні характеристики, включаючи швидкість, прискорення, а також ривки рухів, які необхідні для проектування управління, оптимального в різних сенсах.

Ключові слова: робототехніка; колісні платформи; ривок; керованість; математична модель; комп'ютерне моделювання.

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ УПРАВЛЯЕМОСТИ РОБОТОТИЗИРОВАННЫХ КОЛЕСНЫХ ПЛАТФОРМ С УЧЕТОМ ОГРАНИЧЕНИЙ РЫСКОВ ДВИЖЕНИЙ

Компьютерное моделирование рассматривается как необходимый инструмент для проектирования подходящих автономных систем управления, оптимальных в различных смыслах и, особенно, в обеспечении ограничения рывков движений роботизированных колесных платформ. **Предметом** исследования является разработка теории и методов компьютерного моделирования процессов управляемости роботизированных колесных платформ. **Целью** данного исследования является рассмотрение рывков колесных платформ, причем рывки определяются производными по времени ускорения центра масс колесной платформы, так что эта производная рассматривается как количественная оценка рывков. Обсуждается некорректность по Адамару вычисления рывков прямым дифференцированием фазовых координат в случае компьютерного моделирования с использованием численных методов. **Задачи** настоящего исследования заключаются в разработке обобщенных подходов к математическому и компьютерному моделированию; в теоретическом получении свойств, присущих колесным платформам и пригодных для верификации результатов компьютерного моделирования; в проведении расчетов для получения количественных результатов о процессах управляемости для частного случая электромеханической четырехколесной платформы при прямолинейном движении в режиме разгона из состояния покоя. Методы исследования основаны на уравнениях Лагранжа второго рода, а также на электромеханических аналогиях и окончательном представлении математических моделей в виде системы обыкновенных дифференциальных уравнений первого порядка с начальными условиями для дальнейшего численного решения. Компьютерное моделирование выполняется с использованием бесплатного программного обеспечения Scilab с открытым исходным кодом. **Результаты** исследования заключаются в том, что предложен корректный способ вычисления рывков по фазовым координатам без их дифференцирования, что позволяет исключить некорректность в смысле Адамара, а также в изучении процессов управляемости электромеханической колесной платформы, в том числе результатов для скоростей, ускорений и рывков, необходимых для иллюстрации процессов управляемости роботизированных колесных платформ. Путем сравнения с теоретически установленными собственными свойствами колесных платформ показана правильность результатов компьютерного моделирования. **Выводы** по данному исследованию заключаются в том, что разработанные подходы к компьютерному моделированию процессов управляемости роботизированных колесных платформ позволяют учитывать влияние управления на различные характеристики, включая скорость, ускорение, а также рывки движений, необходимые для проектирования оптимального в разных смыслах управления.

Ключевые слова: робототехника; колесные платформы; рывок; управляемость; математическая модель; компьютерное моделирование.

Бібліографічні описи / Bibliographic descriptions

Альохіна С. В., Невлюдов І. Ш., Ромашов Ю. В. Комп'ютерне моделювання процесів керованості для роботизованих колісних платформ з врахуванням обмежень ривків рухів. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 65–75. DOI: <https://doi.org/10.30837/ITSSI.2022.19.065>

Alyokhina, S., Nevliudov, I., Romashov, Y. (2022), "Computer simulations of controllability processes for robotic wheeled platforms taking into account restrictions of jerk motions", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 65–75. DOI: <https://doi.org/10.30837/ITSSI.2022.19.065>

О. М. ЧЕРНЯК, Н. А. СОРОКОЛАТ, І. О. БАГАЄВ, Л. Ю. ФАТЄЄВА

ЗАСТОСУВАННЯ ФУНКЦІОНАЛЬНОЇ ЗАЛЕЖНОСТІ ДЛЯ БАГАТОКРИТЕРІАЛЬНОГО ОЦІНЮВАННЯ БЕЗПЕКИ ПРАЦІ, ЯК ОБ'ЄКТА КВАЛІМЕТРІЇ

Предметом дослідження в статті є методика багатокритеріального оцінювання показника безпеки праці на виробництві. **Мета** роботи – розробка методики застосування функціональної залежності для багатокритеріального оцінювання безпеки праці на виробництві з урахуванням особливостей системи шкідливих та небезпечних виробничих чинників. В статті вирішується наступне **завдання**: дослідити можливість застосування функціональної залежності для отримання оцінок різномірних показників безпеки праці на безрозмірній шкалі, при цьому врахувати вид її нерівномірності та узгодити її з особливостями небезпечних чинників на виробництві. Використовуються **методи**: кваліметрії та математичної статистики; експертні. Отримано наступні **результати**: проаналізовано існуючі функціональні залежності між вимірними значеннями показників якості та їх оцінкою на безрозмірній шкалі, які застосовували для оцінювання об'єктів кваліметрії різної природи. Показано, що для об'єктивного оцінювання об'єктів кваліметрії, як правило, потрібно застосовувати нелінійні залежності. Головним завданням дослідника вибрати вид нелінійної залежності, це потребує додаткових наукових досліджень. Інструментом для вибору тої чи іншої нелінійної залежності є розуміння фізичної суті об'єкту кваліметрії, тобто розуміння закономірностей зв'язку між вимірним значенням показників кваліметрії та їх оцінкою. Для цього важливо застосування методів експертних оцінок, адже, як правило, такі закономірності невідомі. Функціональна залежність, яка застосовується для отримання оцінок показників безпеки праці на виробництві є ступеневою і включає параметр форми. Змінюючи параметр форми, змінюється кривизна залежності, цим самим змінюється оцінка на безрозмірній шкалі. Така особливість застосованої залежності дає можливість розробити універсальну методику, тобто, змінюючи параметр форми, дану залежність можна застосовувати для різних показників безпеки праці на будь-якому виробництві. Як приклад, у статті розглядається металургійне виробництво, при цьому оцінюються найбільш небезпечні чинники. Розроблено покрокову методику визначення узагальненого показника безпеки праці на виробництві та на прикладі вимірних чисельних значень небезпечних чинників, показано її дієвість та універсальність. **Висновки**: Для визначення узагальненого показника безпеки праці на виробництві пропонується застосування функціональної залежності між окремими показниками шкідливих виробничих чинників та їх значеннями на безрозмірній шкалі, це дає кількісно оцінити безпеку праці на виробництві. За допомогою запропонованої методики можна приймати управлінські рішення, що призводить до мінімізації відхилення дійсних значень шкідливих чинників від оптимальних.

Ключові слова: функціональна залежність; методика оцінювання; багатокритеріальне оцінювання; об'єкт кваліметрії; узагальнений показник; безпека праці.

Вступ

Для успішного розвитку виробничого підприємства, забезпечення конкурентоспроможності продукції та досягнення поставлених стратегічних цілей недостатньо лише вдосконалювати технології, а необхідно забезпечити безпечні умови праці. Для комплексної підтримки безпеки впроваджується система управління безпекою праці, яка враховує процеси ідентифікації, контролю, оцінки та управління. Для ефективності системи контролю необхідно мати науково обґрунтовану методику кількісної оцінки шкідливих факторів з метою контролю та мінімізації їх впливу на організм людини з урахуванням особливостей технологічних процесів.

Технологічні процеси характеризуються: важкістю та інтенсивністю виконуваної роботи; ергономічністю робочого місця; технологічна безпека; рівень технологічної дисципліни; загальна організація робочого процесу; шкідливі та небезпечні фактори. Для комплексної оцінки стану безпеки праці необхідно мати кількісні показники шкідливих виробничих чинників з урахуванням їх технологічних особливостей.

Для отримання кількісних показників якості об'єктів різної природи використовуються методи кваліметрії. Кваліметрія - предмет науки, що вивчає методологію кількісної оцінки якості об'єктів і

процесів різної природи. Під об'єктом кваліметрії в даній статті ми розглядаємо систему шкідливих виробничих чинників які мають вплив на здоров'я та життя працівників.

Система шкідливих виробничих чинників, як об'єкт кваліметрії, має ряд особливостей: можливість проведення вимірювань; різноманітні одиниці та діапазон вимірювань; різний ступінь впливу на організм людини; різна тяжкість наслідків захворювання; можливість використання спеціальних засобів захисту. Такі особливості зумовлюють складність вирішення проблеми кількісної оцінки безпеки праці на виробництві. Для вирішення такої задачі необхідно нанести на безрозмірну шкалу різномасштабні показники різних факторів, що дасть змогу кількісно оцінити узагальнений показник стану безпеки праці в різних галузях економіки.

Аналіз останніх досліджень і публікацій

На працівника в процесі трудової діяльності впливають шкідливі та небезпечні чинники, внаслідок дії яких погіршується його здоров'я та може привести до непрацездатності. Той факт, що шкідливі та небезпечні чинники мають різні одиниці та діапазон вимірювання ускладнює зробити висновок про комплексну кількісну оцінку стану безпеки праці на робочому місці.

Для переведення розмірностей шкідливих та небезпечних факторів у безрозмірну шкалу пропонується застосування нелінійних математичних залежностей між вимірними показниками шкідливих чинників та їх оцінками на безрозмірній шкалі. Як математичні залежності пропонується застосовувати відомі залежності, виведені Гнеденко і раніше застосовані для оцінювання надійності технічних систем, і мають відношення до теорії екстремальних статистик. Використовуючи теорію екстремальних статистик та запропоновані математичні залежності Гнеденко, дослідник Харрінгтон запропонував методику багатокритеріального оцінювання житлових приміщень [1].

У роботах [2-3] для отримання оцінок різнорозмірних показників якості на безрозмірній шкалі використовували залежність, яка мала подвійний експонентний вигляд. Існує ряд робіт, які пов'язані з оцінюванням якості об'єктів різної природи [4-6], в яких досліджувались інструменти для забезпечення єдності оцінювання якості різних об'єктів. Авторами [7-9] використовувався вид залежностей для тримання оцінки показників якості на безрозмірній шкалі, з застосуванням порядкових статистик. В наукових роботах вчених [10-12] застосовуються багатокритеріальні методи кількісної оцінки процесів, які дозволяють привести до однієї розмірності їх показники. У наукових роботах [13, 14] були запропоновані методики оцінювання якості об'єктів кваліметрії, застосовуючи інтегрування методом трапецій та середніх прямокутників.

Метою роботи є розроблення методики застосування функціональної залежності для багатокритеріального оцінювання безпеки праці на виробництві з урахуванням особливостей системи шкідливих та небезпечних виробничих чинників.

Вирішення завдання

Проаналізовані методики оцінювання застосовуються для оцінки якості продукції і на різні процеси системи менеджменту якості, тому вважаємо, що проведений огляд літератури підтвердив актуальність дослідження, спрямоване на розвиток методів оцінювання безпеки праці.

Пропонується застосовувати вид залежностей, який дозволить оцінити стан безпеки праці на виробництві [15]:

$$f(q) = \frac{1}{1 + ab^{-kq}}, \quad (1)$$

де коефіцієнти a та b знаходяться як:

$$b = b_1^{\frac{1}{(q_{\min} - q_{\max})k}}, \quad (2)$$

де:

$$b_1 = \frac{(1 - q_{\max})q_{\min}}{(1 - q_{\min})q_{\max}}, \quad (3)$$

q_{\min} , та q_{\max} – мінімально-допустиме та максимально-допустиме значення показника шкідливого чинника.

Таким чином, після нормування, коефіцієнт a знаходиться як:

$$a = \frac{(1 - q_{\min})}{q_{\min}} b^{kq_{\min}}. \quad (4)$$

Функція (1) має точку перегину при: $q_{\text{пер}} = \frac{\ln a}{k \ln b}$.

Параметр k впливає на крутизну функції вздовж осі ОХ. Змінюючи k , можна керувати кривизною функції (1) і таким чином отримувати різні оцінки при однакових результатах вимірів показника шкідливого чинника.

В графічному вигляді залежність (1) при $k = 1$, показана на рисунку 1.

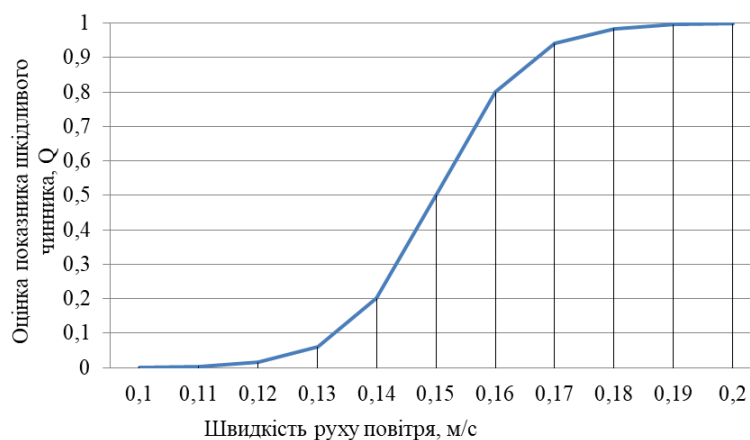


Рис. 1. Вид залежності (1) при $k = 1$

На рисунку 1 показник швидкості руху повітря, який змінюється від $S_{\min} = 0,1$ м/с до $S_{\max} = 0,2$ м/с з

кроком 0,01, щоб наочно показати вид залежності (1). Не залежно від одиниць виміру та від розрядів значення показника шкідливого чинника на

осі ОХ, вид і форма залежності не будуть змінюватися. До зміни форми приведе зміна параметра форми – k .

На рисунку 2 показано серію із п'яти залежностей, при чому коефіцієнт k міняється від 1,2 до 0,8 з кроком 0,1 з ліва на право. По осі ОХ шкала відповідає шкалі одиниць вимірювання і може бути змінена для кожного окремого показника шкідливого чинника. Крайнє ліве значення на цій шкалі (0) відповідає Q_{min} – мінімально-допустимому значенню

показника шкідливого чинника. Крайнє праве значення на цій шкалі (1) відповідає Q_{max} – максимально-допустимому значенню показника шкідливого чинника. Вони відповідають відповідним значенням кожного окремо показника шкідливого чинника. Отже, якщо умови праці мають n показників шкідливих чинників, то для кожного з них необхідно визначити коефіцієнти a та b за формулами (4) та (2) відповідно.

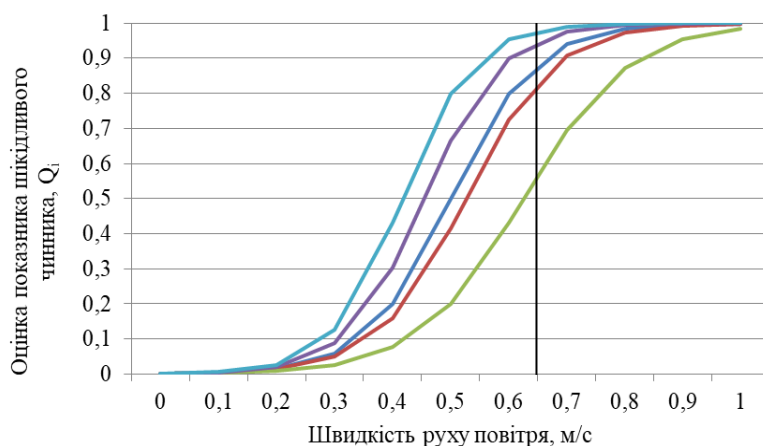


Рис. 2. Серія залежностей (1) при $k = 1,2 - 0,8$ з кроком 0,1 з ліва на право

З рисунка 2 видно, що при показнику якості 0,7 м/с, вимірюваному в одиницях вимірювання, його оцінка на безрозмірній шкалі змінюється від 0,98 при $k = 1,2$ до 0,55 при $k = 0,8$. Отже маючи одне дійсне значення показника шкідливого чинника можемо отримати діапазон значень його оцінок на безрозмірній шкалі. Таким чином можна вибирати один із показників ступені у функції (1) та змінювати оцінку дійсного показника шкідливого чинника на безрозмірній шкалі.

Так як оцінки одиничних показників шкідливих чинників мають однакову шкалу вимірювання (0 - 1), то можна знайти узагальнений показник безпеки праці, застосувавши одну із середніх значень (арифметична, геометрична, гармонійна), які дають змогу звести окремі оцінки. Визначення середнього геометричного значення дасть оцінку безпеки праці нуль, якщо один із показників дорівнюватиме нулю, а одиницю, тобто максимальне значення, можна отримати тільки у випадку, коли всі одиничні показники рівні одиниці. У такому разі узагальнений показник шкідливого чинника розраховується за формулою:

$$W = \sqrt[n]{\prod_{i=1}^n Q_i}, \quad (5)$$

де n – кількість одиничних показників шкідливого чинника; Q_i – значення i -го одиничного показника шкідливого чинника на безрозмірній шкалі.

Таким чином, отримане значення узагальненого показника шкідливих факторів дає підстави приймати

рішення щодо подальших дій щодо покращення безпеки праці.

Пропонуємо покрокову методику визначення узагальненого показника безпеки праці:

Крок 1. Визначити перелік шкідливих та небезпечних чинників на виробництві.

Крок 2. Для кожного шкідливого та небезпечного чинника за нормативними документами визначити максимально - допустиме, мінімально – допустиме та оптимальне (найкраще) значення.

Крок 3. Провести вимірювання цих шкідливих чинників, і результати занести в спеціально підготовлену таблицю.

Крок 4. За формулою (1) визначити оцінку показника шкідливого чинника на безрозмірній шкалі. Для цього необхідно використовувати формули (2 – 4) та призначити параметр форми k . Параметр форми призначається групою експертів.

Крок 5. Визначити узагальнений показник безпеки праці використовуючи формулу (5) з огляду на всі поодинокі показники.

Для підтвердження працездатності методики з оцінювання безпеки праці було проведено дослідження на машинобудівному підприємстві. Для оцінювання розглядалися шкідливі виробничі чинники у ливарному цеху. Ливарний цех обраний як приклад, так як це виробництво з присутністю шкідливих та небезпечних чинників які впливають для здоров'я людини. Визначили, що у ливарному цеху основними шкідливими виробничими чинниками є: мікроклімат (температура повітря, відносна вологість повітря, швидкість руху повітря, інтенсивність теплового випромінювання), шум, локальна вібрація.

Значення шкідливих виробничих факторів вимірювалися та фіксувалися на робочих місцях та у робочій зоні протягом місяця (31 день). Для вимірювання температури повітря, відносної вологості повітря та швидкості руху повітря було використано пристрій – комбінований прилад FLIR EM54. Інтенсивність теплового випромінювання вимірювалася радіометром теплового випромінювання "ІК-метр". Вимірювання рівня шуму

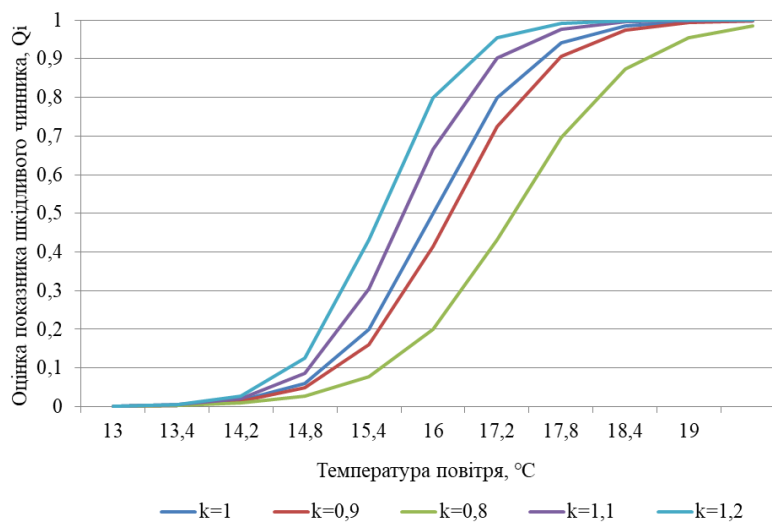
та загальної вібрації здійснювалося цифровим вимірником рівня звуку GM1351 та віброметром AR63A (GM63A). Допустимі норми шкідливих чинників визначено на підприємстві відповідно до чинних нормативних документів.

Отримані експериментальні значення наведених вище показників шкідливих чинників та результати математичних перетворень наведено у таблиці 1.

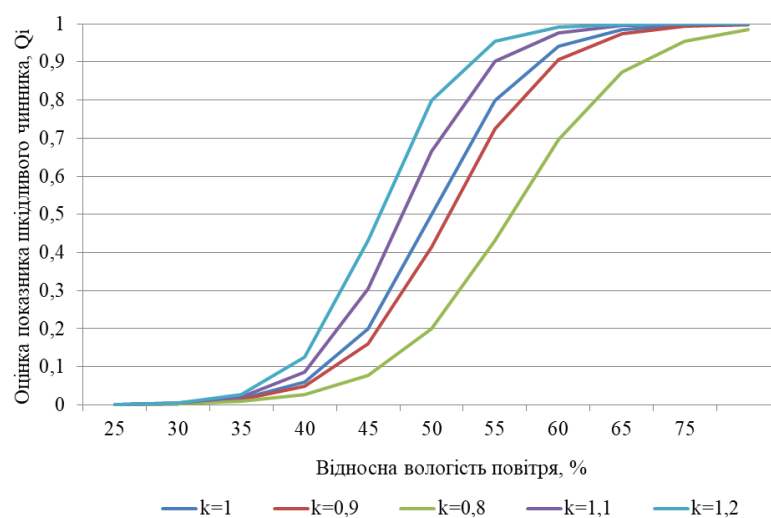
Таблиця 1. Результати впровадження методики оцінки безпеки праці на виробництві

№	Показники шкідливих чинників	q_{min}	q_{max}	q_{omn}	q_i	k	Q
1	Температура повітря, °C	13	19	16	17	1	0,99
2	Відносна вологість повітря, %	25	75	50	70	0,8	0,5
3	Швидкість руху повітря, м/с	0	0,5	0	0,1	1,1	0,9
4	Інтенсивність теплового випромінювання, Вт/м ²	0	140	0	94	1	0,75
5	Шум, дБа	60	80	70	65	1	0,5
6	Локальна вібрація, м/с ²	0	0,2	0	0,09	1,2	0,88

На рисунку 3 представлено графічні зображення залежностей для оцінювання шкідливих виробничих чинників.

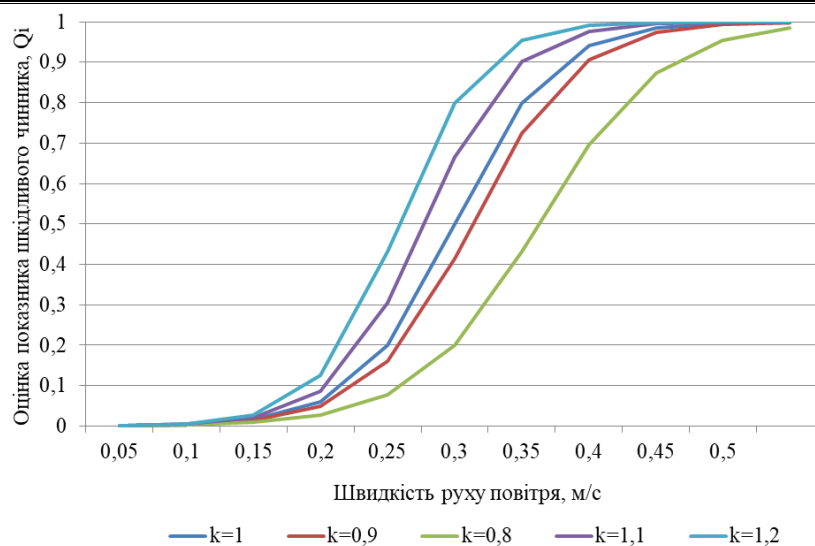


а)

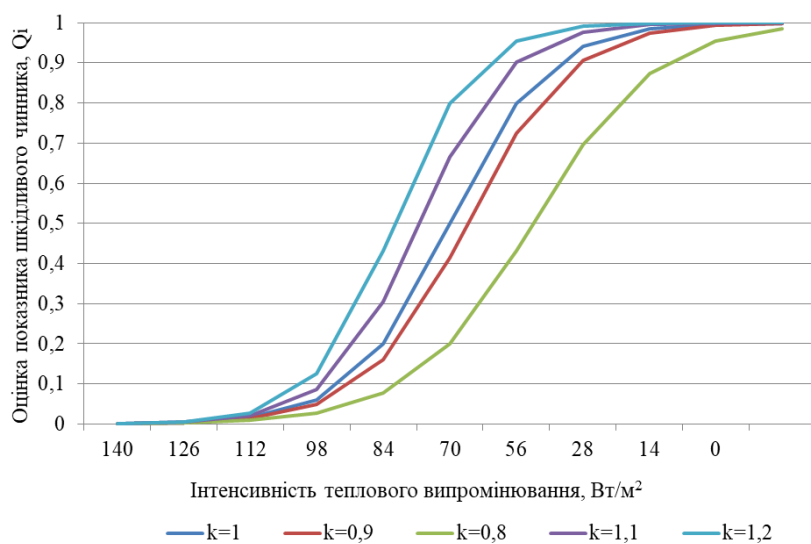


б)

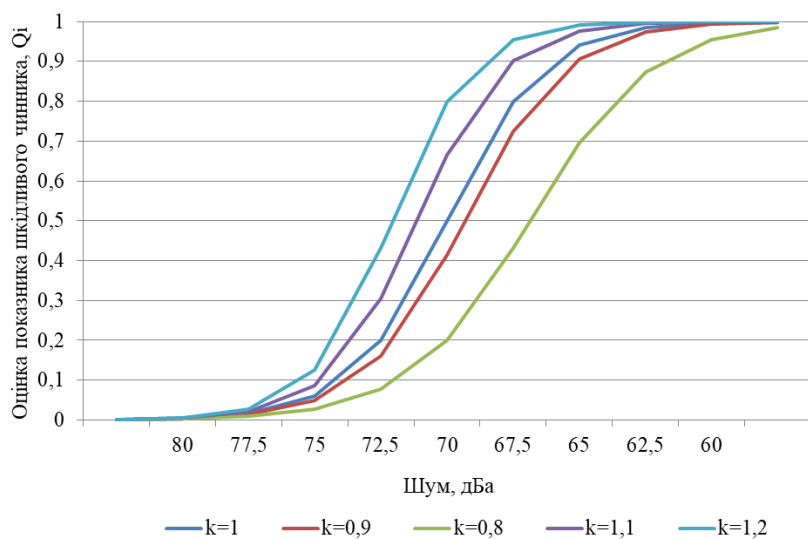
Рис. 3. Графічне зображення залежностей для оцінювання показників шкідливих чинників: а) температури повітря; б) відносної вологості повітря



в)



з)



д)

Рис. 3. Графічне зображення залежностей для оцінювання показників шкідливих чинників: в) швидкості руху повітря; г) інтенсивності теплового випромінювання; д) шуму;

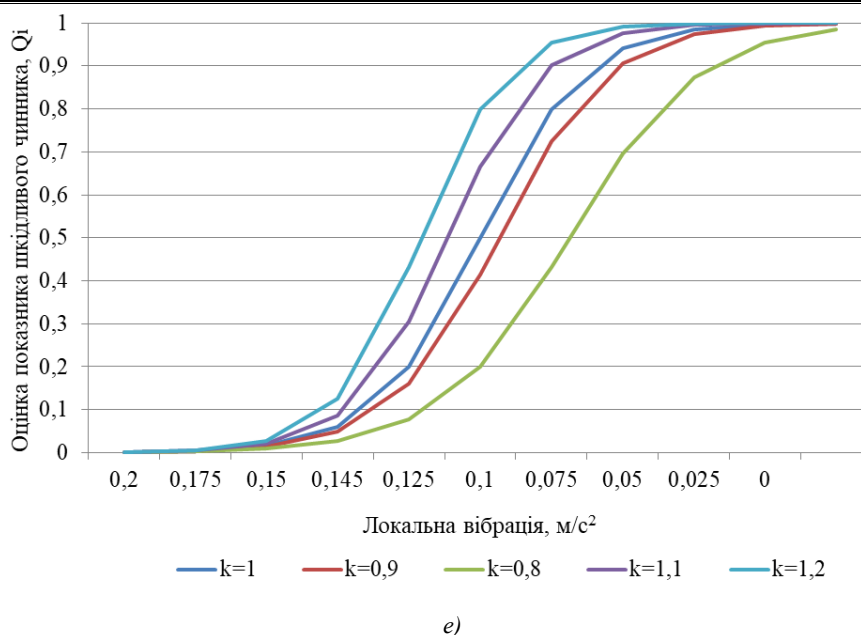


Рис. 3. Графічне зображення залежностей для оцінювання показників шкідливих чинників: е) локальної вібрації.

Знайдемо узагальнений показник, застосувавши одну із середніх значень. В даному випадку застосовується середнє геометричне значення.

$$W = \sqrt[5]{0,99 \cdot 0,5 \cdot 0,9 \cdot 0,75 \cdot 0,5 \cdot 0,88} = 0,73.$$

Отже, за допомогою функціональної залежності та експертів, можна отримувати узагальнений показник безпеки праці на виробництві. Розроблена методика оцінювання являється універсальною, та може бути застосована для оцінювання безпеки праці на виробництвах різних галузей національної економіки.

Висновки

Для оцінювання безпеки праці застосовуються інструменти кваліметрії, оскільки дозволяють

отримати кількісну оцінку шкідливих та небезпечних чинників. Визначено, що система шкідливих виробничих чинників як об'єкт кваліметрії має ряд особливостей, які необхідно враховувати при вирішенні завдання кількісного оцінювання безпеки праці на виробництві.

Таким чином, застосування функціональної залежності між окремими показниками шкідливих виробничих чинників та їх значеннями на безрозмірній шкалі забезпечує отримання кількісної оцінки безпеки праці на виробництві. За допомогою запропонованої методики можна приймати управлінські рішення, що призводять до мінімізації відхилення дійсних значень шкідливих чинників від оптимальних.

Список літератури

1. Тріш Г. М. Розробка методології оцінювання процесів систем управління якістю підприємств з урахуванням вимог міжнародних стандартів: дис. канд. техн. наук : 05.01.02 "Стандартизація, сертифікація та метрологічне забезпечення". 2014. Харків, 162 с.
2. Ginevicius R., Trishch H., Petraskevicius V. Quantitative assessment of quality management systems' processes. *Economic Research-Ekonomska Istraživanja*. 2015. Vol. 1. №. 28. P. 1096–1110. DOI: <https://doi.org/10.1080/1331677X.2015.1087676>.
3. Тріш Г. М., Денисенко М. В. Аналіз стану системи управління якістю в динаміці. *Технологический аудит и резервы производства*. 2014. №1/5 (15). С. 14–16. DOI: <https://doi.org/10.15587/2312-8372.2014.21717>.
4. Trishch R., Cherniak O., Kupriyanov O., Luniachek V., Tsykhanovska I. Methodology for multi-criteria assessment of working conditions as an object of qualimetry. *Engineering Management in Production and Services*. 2021. №13 (2). P. 107–114. DOI: <https://doi.org/10.2478/emj-2021-0016>.
5. Trishch R., Nechuiviter O., Dyadyura K., Vasilevskyi O., Tsykhanovska I., Yakovlev M. Qualimetric method of assessing risks of low quality products. *MM Science Journal*. 2021. №4. P.4769–4774. DOI: https://doi.org/10.17973/MMSJ.2021_10_2021030.
6. Argotti Y., Baron C., Esteban P. Quality quantification in Systems Engineering from the Qualimetry Eye. *2019 IEEE International Systems Conference (SysCon)*. 2019. P. 1–8. DOI: <https://doi.org/10.1109/SYSCON.2019.8836756>.
7. Горбенко Н. А., Катрич О. А. Оценивание процессов системы менеджмента качества предприятий на соответствие требований международных стандартов серии ISO 9000. *Вісник Національного технічного університету „ХПІ”*: Збірник наукових праць. Серія: Механіко-технологічні системи та комплекси. 2014. № 40 (1083). С. 20–25.

8. Черкашина О. С. Керування якістю технологічних процесів в машинобудуванні з застосуванням трипараметричного моделювання. *Машинобудування*. 2019. №23. С. 159–165. DOI: <https://doi.org/10.32820/2079-1747-2019-23-159-165>.
9. Ким Н. И., Трищ Р. М. Закономерности рассеивания безразмерных показателей качества объектов различной природы. *Системы управління, навігації та зв'язку. Серія: Метрологія та вимірювальна техніка*. 2016. № 4 (40). С. 143–145.
10. Zhang C., Chen C., Streimikiene D., Balezentis T. Intuitionistic fuzzy MULTIMOORA approach for multi-criteria assessment of the energy storage technologies. *Applied Soft Computing*. 2019. №79. P. 410–423. DOI: <https://doi.org/10.1016/j.asoc.2019.04.008>.
11. Alamerew Y.A., Kambanou M. L., Sakao T., Brissaud D. A Multi-Criteria Evaluation Method of Product-Level Circularity Strategies. *Sustainability*. 2020. 12 (12): 5129. DOI: <https://doi.org/10.3390/su12125129>.
12. Zolfani S., Maknoon R., Zavadskas E. Multiple Nash Equilibriums and evaluation of strategies. New application of MCDM methods. *Journal of Business Economics and Management*. 2015. № 16. P. 290–306. DOI: <https://doi.org/10.3846/16111699.2014.967715>.
13. Черняк О. М., Сороколат Н. А., Каницька І. В. Застосування методу інтегрування для оцінювання якості об'єктів кваліметрії. *Вісник Національного технічного університету «ХПІ»*. Серія: Нові рішення в сучасних технологіях. Харків: НТУ «ХПІ». 2020. № 4 (6). С. 93–98. DOI: <https://doi.org/10.30837/ITSSI.2020.14.169>.
14. Черняк О. М., Сороколат Н. А., Каницька І. В. Графоаналітичний метод визначення комплексного показника якості об'єктів кваліметрії. *Сучасний стан наукових досліджень та технологій в промисловості*. 2020. № 4 (14). С. 169–175. DOI: <https://doi.org/10.30837/ITSSI.2020.14.169>.
15. Ким Н. И. Узагальнений показник якості об'єктів кваліметрії різної природи. *Вісник аграрної науки Причорномор'я*. 2021. Вип. 1. С. 94–101. DOI: [https://doi.org/10.31521/2313-092X/2021-1\(109\)-12](https://doi.org/10.31521/2313-092X/2021-1(109)-12).

References

1. Trishch, G. M. (2014), "Development of a methodology for evaluating the processes of enterprise quality management systems taking into account the requirements of international standards" ["Rozrobka metodolohiyi otsynuvannya protsesiv system upravlinnya yakistyu pidpriyemstv z urakhuvannyam vymoh mizhnarodnykh standartiv"], *dissertation of candidate of technical sciences tech. Sciences*: 05.01.02 "Standardization, certification and metrological support", Kharkiv, 162 p.
2. Ginevičius, R., Trishch, H., Petraškevičius, V. (2015), "Quantitative assessment of quality management systems' processes", *Economic Research-Ekonomska Istraživanja*, No. 28:1, P. 1096–1110. DOI: <https://doi.org/10.1080/1331677X.2015.1087676>.
3. Trishch, G. M., Denysenko, M. V. (2014), "Analysis of Quality Management System in the Dynamics" ["Analiz stanu systemy upravlinnya yakistyu v dynamitsi"], *Technology Audit and Production Reserves*, No 1 (5), P. 14–16, DOI: <https://doi.org/10.15587/2312-8372.2014.21717>.
4. Trishch, R., Cherniak, O., Kupriyanov, O., Luniachek, V., Tsykhanovska, I. (2021), "Methodology for multi-criteria assessment of working conditions as an object of qualimetry", *Engineering Management in Production and Services*, No. 13(2), P. 107–114. DOI: <https://doi.org/10.2478/emj-2021-001612>.
5. Trishch, R., Nechuviter, O., Dyadyura, K., Vasilevskyi, O., Tsykhanovska, I., Yakovlev, M. (2021), "Qualimetric method of assessing risks of low quality products", *MM Science Journal*, No 4, P. 4769–4774. DOI: https://doi.org/10.17973/MMSJ.2021_10_2021030.
6. Argotti, Y., Baron, C., Esteban, P. (2019), "Quality quantification in Systems Engineering from the Qualimetry Eye", *IEEE International Systems Conference (SysCon)*, P. 1–8. DOI: <https://doi.org/10.1109/SYSCON.2019.8836756>.
7. Gorbenko, N. A., Katrich, O. A. (2014), "Evaluation of the processes of the quality management system of enterprises for compliance with the requirements of international standards of the series ISO 9000" ["Otsenivaniye protsessov systemy menedzhmenta kachestva predpriyatiy na sootvetstviye trebovaniy mezhdunarodnykh standartov serii ISO 9000"], *Bulletin of the National Technical University "KhPI": Collection of scientific papers. Series: Mechanical-technological systems and complexes*, No. 40 (1083), P. 20–25.
8. Cherkashina, O. (2019), "Management of quality of technological processes in mechanical engineering using three-parameter modeling" ["Keruvannya yakistyu tekhnolohichnykh protsesiv v mashynobuduvanni z zastosuvannyam tryparametrychnoho modelyuvannya"], *Engineering*. No. 23, P. 159–165. DOI: <https://doi.org/10.32820/2079-1747-2019-23-159-165>.
9. Kim, N.I., Trishch, R.M. (2016), "Laws of dispersion quality indicators dimensionless objects of different nature" ["Zakonomernosti rasseivaniya bezrazmernykh pokazateley kachestva ob'yektiv razlichnoy prirody"], *Control, navigation and communication systems. Series: Metrology and measuring technology*, No. 4 (40), P. 143–145.
10. Zhang, C., Chen, C., Streimikiene, D., Balezentis, T. (2019), "Intuitionistic fuzzy MULTIMOORA approach for multi-criteria assessment of the energy storage technologies", *Applied Soft Computing*, No. 79, P. 410–423. DOI: <https://doi.org/10.1016/j.asoc.2019.04.008>.
11. Alamerew, Y.A.; Kambanou, M.L.; Sakao, T., Brissaud, D. (2020), "A Multi-Criteria Evaluation Method of Product-Level Circularity Strategies", *Sustainability*, No. 12, 5129. DOI: <https://doi.org/10.3390/su12125129>.
12. Zolfani, S., Maknoon, R., Zavadskas, E. (2015), "Multiple Nash Equilibriums and evaluation of strategies. New application of MCDM methods", *Journal of Business Economics and Management*, No. 16, P. 290–306. DOI: <https://doi.org/10.3846/16111699.2014.967715>.
13. Cherniak, O., Sorocolat, N., Kanytska, I. (2020), "Application of the integration method for assessing the quality of qualimetry objects" ["Zastosuvannya metodu intehruvannya dlya otsynuvannya yakosti ob'yektiv kvalimetriyi"], *Bulletin of*

- the National Technical University "KhPI". Series: New solutions in modern technology, Kharkiv, NTU "KhPI", No. 4 (6), P. 93–98. DOI: <https://doi.org/10.20998/2413-4295.2020.04.14>.
14. Cherniak, O., Sorocolat, N., Kanytska, I. (2020), "Graph analytical method for determining the complex quality indicator of qualimetry objects", ["Hrafoanalytychnyy metod vyznachennya kompleksnoho pokaznyka yakosti ob'ektiv kvalimetriyi"], *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (14), P. 169–175. DOI: <https://doi.org/10.30837/ITSSI.2020.14.169>.
 15. Kim, N. I. (2021), "Generalized indicator of qualimetry objects quality of various nature" ["Uzahal'neny pokaznyk yakosti ob'ektiv kvalimetriyi riznoyi pryrody"], *Bulletin of Agrarian Science of the Black Sea Coast*, No. 1, P. 94–101. DOI: [https://doi.org/10.31521/2313-092X/2021-1\(109\)-12](https://doi.org/10.31521/2313-092X/2021-1(109)-12).

Надійшла (Received) 29.03.2022

Відомості про авторів / Сведения об авторах / About the Authors

Черняк Олена Миколаївна – кандидат технічних наук, доцент, Українська інженерно-педагогічна академія, старший викладач кафедри автоматизації, метрології та енергоефективних технологій, Харків, Україна; e-mail: olena-cherniak@ukr.net; ORCID: <https://orcid.org/0000-0001-6167-8809>.

Черняк Елена Николаевна – кандидат технических наук, доцент, Украинская инженерно-педагогическая академия, старший преподаватель кафедры автоматизации, метрологии и энергоэффективных технологий, Харьков, Украина.

Cherniak Olena – PhD (Engineering Sciences), Associate Professor, Ukrainian Engineering Pedagogics Academy, Senior Lecturer of the Department of Automation, Metrology and Energy Efficient Technologies, Kharkiv, Ukraine.

Сороколат Наталія Андріївна – Українська інженерно-педагогічна академія, аспірантка кафедри автоматизації, метрології та енергоефективних технологій; Харків, Україна; e-mail: n.a.sorokolat@gmail.com; ORCID: <https://orcid.org/0000-0002-0140-9364>.

Сороколат Наталия Андреевна – Украинская инженерно-педагогическая академия, аспирантка кафедры автоматизации, метрологии и энергоэффективных технологий, Харьков, Украина.

Sorocolat Nataliia – Ukrainian Engineering Pedagogics Academy, graduate student Department of Automation, Metrology and Energy Efficient Technologies, Kharkiv, Ukraine.

Багаєв Ігор Олександрович – Українська інженерно-педагогічна академія, аспірант кафедри автоматизації, метрології та енергоефективних технологій; Харків, Україна; e-mail: i.a.bagayev@gmail.com; ORCID: <https://orcid.org/0000-0002-9101-5114>.

Багаев Игорь Александрович – Украинская инженерно-педагогическая академия, аспирант кафедры автоматизации, метрологии и энергоэффективных технологий, Харьков, Украина.

Ihor Bahaiev – Ukrainian Engineering Pedagogics Academy, graduate student Department of Automation, Metrology and Energy Efficient Technologies, Kharkiv, Ukraine.

Фатєєва Ліна Юріївна – Українська інженерно-педагогічна академія, аспірантка кафедри автоматизації, метрології та енергоефективних технологій; Харків, Україна; e-mail: linafat81@gmail.com; ORCID: <https://orcid.org/0000-0002-6460-0772>.

Фатеева Лина Юрьевна – Украинская инженерно-педагогическая академия, аспирантка кафедры автоматизации, метрологии и энергоэффективных технологий, Харьков, Украина.

Lina Fatieieva – Ukrainian Engineering Pedagogics Academy, graduate student Department of Automation, Metrology and Energy Efficient Technologies, Kharkiv, Ukraine.

ПРИМЕНЕНИЕ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ ДЛЯ МНОГОКРИТЕРИАЛЬНОГО ОЦЕНИВАНИЯ БЕЗОПАСНОСТИ ТРУДА, КАК ОБЪЕКТА КВАЛИМЕТРИИ

Предметом исследования в статье является методика многокритериального оценивания показателя безопасности труда на производстве. **Цель** работы – разработка методики применения функциональной зависимости для многокритериального оценивания безопасности труда на производстве с учетом особенностей системы вредных и опасных производственных факторов. В статье решается следующая **задача**: исследовать возможность применения функциональной зависимости для получения оценок разноразмерных показателей безопасности труда на безразмерной шкале, при этом учесть ее неравномерность и согласовать ее с особенностями опасных факторов на производстве. Используются методы: квалиметрии и математической статистики; экспертные. Получены следующие **результаты**: проанализированы существующие функциональные зависимости между измеренными значениями показателей качества и их оценкой на безразмерной шкале, которые использовались для оценки объектов квалиметрии различной природы. Показано, что для объективной оценки объектов квалиметрии, как правило, следует применять нелинейные зависимости. Главной задачей исследователя выбрать вид нелинейной зависимости, это требует дополнительных научных исследований. Инструмент для выбора той или иной нелинейной зависимости является понимание физической сущности объекта квалиметрии, то есть понимание закономерностей связи между измеренным значением показателей квалиметрии и их оценкой. Для этого важно применение методов экспертных оценок, ведь обычно такие закономерности неизвестны. Функциональная зависимость, применяемая для получения оценок показателей безопасности труда на производстве, является ступенчатой и включает параметр формы.

Изменяя параметр формы, меняется кривизна зависимости, тем самым меняется оценка на безразмерной шкале. Такая особенность применяемой зависимости позволяет разработать универсальную методику, то есть, изменяя параметр формы, данную зависимость можно применять для различных показателей безопасности труда на любом производстве. В качестве примера в статье рассматривается металлургическое производство, при этом оцениваются наиболее опасные факторы. Разработана пошаговая методика определения обобщенного показателя безопасности труда на производстве и примере измеренных численных значений опасных факторов, показаны ее действенность и универсальность. **Выводы:** Для определения обобщенного показателя безопасности труда на производстве предлагается применение функциональной зависимости между отдельными показателями вредных производственных факторов и их значениями на безразмерной шкале, что дает количественную оценку безопасности труда на производстве. С помощью предлагаемой методики можно принимать управленческие решения, что приводит к минимизации отклонения действительных значений вредных факторов от оптимальных.

Ключевые слова: функциональная зависимость; методика оценки; многокритериальное оценивание; объект квалиметрии; обобщенный показатель; безопасность труда.

APPLICATION OF FUNCTIONAL DEPENDENCE FOR MULTI-CRITERIAL ASSESSMENT OF LABOR SAFETY AS AN OBJECT OF QUALIMETRY

The **subject** matter of the article is the method of multi-criteria assessment of the indicator of labor safety in production. The **goal** of the work is to develop a methodology for applying functional dependence for multi-criteria assessment of labor safety in production, taking into account the characteristics of the system of harmful and dangerous production factors. The following **task** is solved in the article: to explore the possibility of using a functional dependence to obtain estimates of different-dimensional labor safety indicators on a dimensionless scale, while taking into account its unevenness and coordinating it with the characteristics of hazardous factors in production. **Methods** are used: qualimetry and mathematical statistics; expert. The following **results** were obtained: the existing functional dependencies between the measured values of quality indicators and their assessment on a dimensionless scale, which were used to evaluate qualimetry objects of various natures, were analyzed. It is shown that for an objective assessment of qualimetry objects, as a rule, nonlinear dependencies should be used. The main task of the researcher is to choose the type of nonlinear dependence; this requires additional scientific research. The tool for choosing one or another non-linear relationship is the understanding of the physical essence of the qualimetry object, that is, the understanding of the patterns of the relationship between the measured value of the qualimetry indicators and their assessment. For this, it is important to use the methods of expert assessments, because usually such patterns are unknown. The functional dependence used to obtain estimates of labor safety indicators in production is stepwise and includes a form parameter. By changing the shape parameter, the curvature of the dependence changes, thereby changing the estimate on the dimensionless scale. This feature of the applied dependence allows us to develop a universal methodology, that is, by changing the shape parameter, this dependence can be applied to various indicators of labor safety in any production. As an example, the article considers metallurgical production, while assessing the most dangerous factors. A step-by-step methodology for determining a generalized indicator of labor safety in production and an example of measured numerical values of hazardous factors has been developed, its effectiveness and versatility have been shown. **Conclusions:** to determine the generalized indicator of occupational safety at work, it is proposed to use the functional relationship between individual indicators of harmful production factors and their values on a dimensionless scale, which gives a quantitative assessment of occupational safety at work. With the help of the proposed method, you can make management decisions, which minimizes the deviation of the actual values of harmful factors from the optimal ones.

Keywords: functional dependence; assessment methodology; multi-criteria assessment; qualimetry object; generalized indicator; labor safety.

Бібліографічні описи / Bibliographic descriptions

Черняк О. М., Сороколат Н. А., Багаєв І. О., Фатєєва Л. Ю. Застосування функціональної залежності для багатокритеріального оцінювання безпеки праці, як об'єкта кваліметрії. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 76–84. DOI: <https://doi.org/10.30837/ITSSI.2022.19.076>

Cherniak, O., Sorocolat, N., Bahaiev, I., Fatieieva, L. (2022), "Application of functional dependence for multi-criterial assessment of labor safety as an object of qualimetry", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 76–84. DOI: <https://doi.org/10.30837/ITSSI.2022.19.076>

A. ZUEV, A. IVASHKO, D. LUNIN

ESTIMATION OF SOFTWARE COMPLEXITY OF CALCULATION OF AUTOREGRESSION COEFFICIENTS AT DIGITAL SPECTRAL ANALYSIS

The **subject** of research in the article are algorithms for fast calculation of autoregression coefficients in digital spectral analysis and estimation of the number of arithmetic operations required for their implementation. The **aim** of the article – comparative analysis of the speed of different algorithms for calculating the coefficients of autoregression as part of the algorithms of spectral analysis, including analysis of the complexity of their microcontroller implementation. **Tasks** to be solved: selection of spectral analysis methods suitable for diagnostics of technological equipment, analysis of methods for calculating autoregression coefficients and derivation of relations for estimating software complexity of algorithms and calculation of numerical estimates of addition and multiplication for some algorithms, adaptation of developed methods and estimates to microcontrollers. spectrum Applied **methods**: algorithm theory, Fourier transform, natural series, microcontroller programming. The **results** obtained: it is shown that spectral estimation methods based on Yul-Walker equations, which require the calculation of autoregression coefficients, combine sufficient resolution and resistance to interference with acceptable implementation complexity. Estimates of the number of additions and multiplications for the Levinson, Durbin, and Trench algorithms are obtained, and their comparative analysis is performed. The calculation times for microcontroller arithmetic with fixed and floating points were counted upon. **Conclusions**: When constructing spectrum analyzers for the diagnosis of technological equipment, it is advisable to use the Yul-Walker method. A comparison of Levinson, Durbin, and Trench algorithms for calculating autoregression coefficients showed that the Trench method requires a minimum number of additions, and the Durbin method requires a minimum number of multiplications. At microcontroller realization of spectrum analyzers, it is necessary to consider features of the arithmetic used by the controller. The Trench method is the fastest in the case of floating-point arithmetic and small-scale modeling. In other cases, Durbin's method is more effective.

Keywords: spectral analysis; autoregression; Levinson; Durbin; Trench algorithms; computational complexity; computer arithmetic; microcontrollers.

Introduction

Monitoring of spectra of mechanical oscillations and signals of energy objects is widely used in solving some problems in the fields of technical and medical diagnostics. By analyzing the amplitudes and locations of the spectral density peaks, we can conclude about the state of the object under study. In particular, changes in the pattern of mechanical vibrations caused by malfunctions of mechanical systems, including turbines and generators, can be quickly detected. On the other hand, the analysis of the spectrum of energy signals allows to assess the quality of electricity and take measures to improve it.

There are a number of spectral estimation algorithms that differ in resolution, resistance to interference, complexity of software and hardware implementation. At the same time, the development of microprocessor technology has made it possible to create portable devices with built-in microcontrollers that provide operational spectral analysis of technical objects in real time. The spectral analysis algorithms used in such devices must ensure acceptable separation and suppression of noise and at the same time a minimum number of addition and multiplication operations.

Sufficient accuracy in combination with acceptable computational complexity is provided by spectral analysis algorithms based on parametric models, in particular Berg's algorithm, covariance algorithm, MUSIC method. The most widespread due to the combination of satisfactory technical characteristics and ease of implementation was the method of Yul-Walker [1]. Its practical application, however, is somewhat complicated by the lack of estimates of the number of computational operations that may be required for a reasonable choice of a microcontroller. Therefore, it is important to obtain the

relations that determine the computational complexity of both the algorithm for spectral estimation in general and its individual components.

Review of existing methods of spectral analysis and estimates of their labor content

Automated diagnostic systems for technological equipment are increasingly used in the creation of the industrial Internet of Things and eventually in the construction of "smart industries". Rhythmic work of modern automated industries is impossible without timely detection and replacement of failed elements. At the same time, the development of modern computer equipment allows the use of complex but effective diagnostic algorithms.

Thus, [1] solves the problem of classification of aircraft engine operating modes based on neural network technologies in real time, and selects the most efficient network architecture. In the article [2] to solve this problem uses the apparatus of fuzzy sets. In this case, the proposed logical inference mechanism eliminates the use of the product rule base, which ensures the practical independence of the computational procedure from the size of the problem. A feature of the methods used in [1, 2] is the use of complex computational algorithms that require significant computing power.

Many papers recommend the use of digital signal processing in technical diagnostic tasks, including analysis of frequency characteristics. Thus, in [3] a method of increasing the stability of estimates of parameters of energy systems, including spectral characteristics, is proposed. However, the main attention in the article is paid to modifications of algorithms for calculation of autocorrelation functions, while methods of calculation of

autoregression coefficients and estimation of their complexity are practically not considered.

Similar issues are studied in [4, 5], where the stability of energy systems is analyzed by monitoring electromechanical oscillations. As diagnostic information in the work it is offered to consider frequencies and attenuation coefficients of the main harmonics of energy signals. The efficiency of the multi-channel Yule-Walker estimation of a multidimensional autoregressive model for determining the parameters of power systems is investigated, and in [4] the model is introduced, and in [5] its decomposition by eigenvectors is considered. Unfortunately, the methods proposed in the works require a significant amount of calculations and are unsuitable for use at local measuring instruments.

Algorithms for solving systems of linear equations with Hankel matrices [6] and Toeplitz [7], which arise in spectral analysis problems, are considered in [6, 7], however, the methods proposed in these works are not always suitable for solving Yul-Walker equation systems. The Levinson-Durbin algorithm for studying autoregression processes is used in [8]. The authors modified the algorithm, providing on the one hand its recursive execution, on the other - the ability to calculate derivatives and the Hesse matrix. However, there is no comparison of the Levinson-Durbin algorithm with other algorithms, in particular, the Trench algorithm.

Issues of speed of spectral analysis algorithms are considered in [9, 10]. The main idea of the article [9] is to estimate the spectral peaks by the method of maximum likelihood. The method developed by the authors significantly reduces the computational complexity of spectral estimation, providing sufficient resolution. In [10], it is proposed to estimate the frequency spectrum by interpolating the discrete Fourier transform coefficients, which allows to detect spectral peaks and adjacent frequencies. In this case, the performance estimates given in [9, 10] are not tied to a specific technical base, which complicates the design of spectrum analyzers.

A number of works consider the possibility of using a modern element base for signal processing. Thus, in [11] the implementation of the method of measuring the parameters of signals and paths on FPGA with increasing the accuracy of optimal information processing is considered, but the work does not provide data on the speed of the proposed technical solutions. The microcontroller implementation of spectrum analyzers is considered in [12, 13].

The aim of the article is a comparative analysis of the speed of different algorithms for calculating the autoregression coefficients as part of the algorithms of spectral analysis, including the microcontroller implementation of some algorithms.

Theoretical analysis of autoregressive algorithms

A group of spectral analysis algorithms based on signal representation as a result of white noise passing through a digital filter has become widespread. In this case, the spectrum $X(\omega)$ coincides with the frequency response of the filter $H(\omega)$ up to a constant multiplier.

Then, restoring the structure and coefficients of the filter, you can calculate its frequency response and the spectrum of the signal $X(\omega)$. If the supposed filter has HIX – the structure containing only the return branch, i.e. in its transfer characteristic there are only poles

$$H(z) = \frac{1}{1 - b_1 \cdot z^{-1} - b_2 \cdot z^{-2} - \dots - b_{M-1} \cdot z^{-(M-1)} - b_M \cdot z^{-M}}, \quad (1)$$

where M is a model order,

b_i – filter coefficients (autoregression coefficients),

then there is an autoregression model (AR model). Determining the spectrum estimate in the case of using the autoregression model occurs in three stages:

1. Calculation of estimates of autocorrelation functions (ACF),

$$R(k) = \frac{1}{N} \sum_{i=0}^{N-1} x_i \cdot x_{i+k}, \quad (2)$$

where x_i – samples of the discrete signal, N is the number of samples, $R(k)$ is the ACF value.

2. Compilation and solution of the so-called Yul-Walker equations

$$\begin{bmatrix} R(0) & R(-1) & \dots & R(-M) \\ R(1) & R(-0) & \dots & R(-M+1) \\ \dots & \dots & \dots & \dots \\ R(M-1) & R(M-2) & \dots & R(1) \\ R(M) & R(M-1) & \dots & R(0) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -b_1 \\ \dots \\ -b_{M-1} \\ -b_M \end{bmatrix} = \begin{bmatrix} D_\epsilon \\ 0 \\ \dots \\ 0 \\ 0 \end{bmatrix}, \quad (3)$$

where M is a model order,

b_i – filter coefficients (autoregression coefficients)

in (1),

D_ϵ – noise dispersion.

3. Calculation of power spectral density $S(\omega)$,

$$S(\omega) = \frac{D_\epsilon}{\left(1 - \sum_{k=1}^M b_k \cos(\omega k)\right)^2 + \left(\sum_{k=1}^M b_k \sin(\omega k)\right)^2}, \quad (4)$$

and the amplitude spectrum $X(\omega) = \sqrt{S(\omega)}$,

where ω is the dimensionless cyclic frequency,

$S(\omega)$ is the spectral power density,

$X(\omega)$ is the amplitude spectrum.

The most labor content stage of calculations is the solution of the system of Yul-Walker equations (3). The application of the traditional Gaussian method requires approximately $\frac{M^3}{3} + \frac{3}{2}M^2 - \frac{5}{6}M$ arithmetic operations.

However, it should be borne in mind that the matrices of coefficients of systems of linear equations have a special form – they have all the elements located on any diagonal, identical, i.e. $a_{i,j} = a_{i+k,j+k}$. Such matrices are called greenhouse (named after the German mathematician O. Toeplitz), and a number of methods have been developed to solve systems with greenhouse matrices of coefficients

– algorithms of Levinson, Durbin, Trench [14]. The use of such algorithms can significantly reduce the number of operations and the amount of memory when solving systems of linear equations and, thus, simplify the process of finding model coefficients and spectral estimates.

Calculation of the number of arithmetic operations

To solve the system of Yul-Walker equations using the Levinson algorithm, it is necessary to make $(M - 1)$ iterations, where M is the order of the model. At each iteration, the scalar variables α , β , and μ , as well as the auxiliary vectors v and z , are calculated. When calculating the scalar variable β by the formula

$$\beta_k = \beta_{k-1}(1 - \alpha_{k-1}^2)$$

it is necessary to perform two multiplication operations and one addition operation. When calculating the scalar variable μ by the formula

$$\mu_k = \frac{b_{k+1} - r^T \cdot E_k x^*}{\beta_k}$$

k additions and $k + 1$ multiplications are performed, where k is the number of the current iteration. To find the vector by the formula

$$v = x + \mu y^*$$

it is necessary to perform k additions and k multiplications. To calculate the scalar variable α by the formula

$$\alpha_k = -\frac{r_{k+1} + r^{(k)T} \cdot y^*}{\beta_k}$$

k additions and $k + 1$ multiplications are spent, finding the auxiliary vector z by the formula

$$z = y + \alpha y^*$$

k additions and k multiplications are performed. Summarizing all arithmetic operations, we obtain the number of additions and multiplications in the algorithm. For the first part of the algorithm, when k changes in the range $1:M - 1$:

Number of additions $S_{k1} = 1 + k + k = 2k + 1$;

Number of multiplications $P_{k1} = 2 + k + 1 + k = 2k + 3$.

For the second part of the algorithm, when $k = 1:M - 2$:

$S_{k2} = k + k = 2k$;

$P_{k2} = k + 1 + k = 2k + 1$.

Calculate the total number of additions and multiplications for each of the stages, using the formula to calculate the sum of the arithmetic progression.

$$S_1 = \sum_{k=1}^{M-1} S_{k1} = \sum_{k=1}^{M-1} (2k + 1) = M^2 - 1,$$

$$P_1 = \sum_{k=1}^{M-1} P_{k1} = \sum_{k=1}^{M-1} (2k + 3) = M^2 + 2M - 3,$$

$$S_2 = \sum_{k=1}^{M-2} S_{k2} = \sum_{k=1}^{M-2} 2k = M^2 - 3M + 2,$$

$$P_2 = \sum_{k=1}^{M-2} P_{k2} = \sum_{k=1}^{M-2} (2k + 1) = M^2 - 2M.$$

Combining these two sums, we finally get the number of additions:

$$S = S_1 + S_2 = 2M^2 - 3M + 1. \quad (5)$$

And the number of multiplications:

$$P = P_1 + P_2 = 2M^2 - 3. \quad (6)$$

When solving the system of Yul-Walker equations based on the Durbin algorithm at each step, the calculation of two scalar variables α , β and the auxiliary vector z is performed, which can be written by such recursive relations

$$\alpha_k = -\frac{r_{k+1} + r^{(k)T} \cdot y^*}{\beta_k},$$

where $\beta_k = \beta_{k-1}(1 - \alpha_{k-1}^2)$.

The number of iterations of the main cycle in the algorithm is equal to $M - 1$. When calculating the scalar variable β it is necessary to perform two multiplication operations and one addition operation. When calculating the scalar variable α , k additions and $k + 1$ multiplications are performed, where k is the number of the current iteration. To calculate the auxiliary vector z by the formula:

$$z_i = y_i + \alpha y_{k+1-i}, \text{ where } i = 1:k.$$

k additions and k multiplications are performed (k is the number of the current iteration). Summarizing the arithmetic operations for one iteration, we obtain: the number of additions at each step $S_k = 1 + k + k = 2k + 1$; the number of multiplications $P_k = 2 + k + 1 + k = 2k + 3$.

Since k depends on the number of iterations, and the number of required iterations is $M - 1$, the total number of additions and multiplications is equal to:

$$S = \sum_{k=1}^{M-1} S_k = \sum_{k=1}^{M-1} (2k + 1) = M^2 - 1, \quad (7)$$

$$P = \sum_{k=1}^{M-1} P_k = \sum_{k=1}^{M-1} (2k + 3) = M^2 + 2M - 3. \quad (8)$$

To solve the system of Yul-Walker equations, using the Trench algorithm, the first step is to calculate the scalar variable γ , as well as the auxiliary vector v by formulas:

$$\gamma = \frac{1}{1 + r^T y^*},$$

$$v = \gamma y^*.$$

Counting the number of transactions, we obtain:

Number of additions $S_1 = M - 1$;

Number of multiplications
 $P_1 = M + M - 1 = 2M - 1$.

In the second stage, a nested cycle is used, which calculates the matrix $B = T^{-1}$. The number of iterations of the inner cycle depends on the current step of iteration of the outer cycle and is equal to: for the outer cycle $i = 2:((M-1)/2 + 1)$, for the inner cycle $j = i:(M-i+1)$. When calculating the elements of the matrix B by the formula:

$$b_{i,j} = b_{n-j,n-i} + \frac{v_i v_j - v_{n-j} v_{n-i}}{\gamma}$$

it is necessary to perform two addition operations and three multiplication operations at each stage of the internal and external cycles. Thus, the number of additions and multiplications in the second stage will be equal:

$$S_2 = \sum_{k=2}^{\frac{M-1}{2}+1} \sum_{i=2}^{M-i+1} S_{k2} = \sum_{k=2}^{\frac{M-1}{2}+1} \sum_{i=2}^{M-i+1} 2 = 2 \cdot \sum_{k=2}^{\frac{M-1}{2}+1} (M-2k+2) = 2 \cdot \frac{M^2 - 2M + 1}{2} = M^2 - 2M + 1$$

$$P_2 = \sum_{k=2}^{\frac{M-1}{2}+1} \sum_{i=2}^{M-i+1} P_{k2} = \sum_{k=2}^{\frac{M-1}{2}+1} \sum_{i=2}^{M-i+1} 3 = 3 \cdot \sum_{k=2}^{\frac{M-1}{2}+1} (M-2k+2) = 3 \cdot \frac{M^2 - 2M + 1}{2} = \frac{3}{2} \cdot (M^2 - 2M + 1).$$

Combining the two sums, we finally get the number of additions and multiplications:

$$S = S_1 + S_2 = M - 1 + M^2 - 2M + 1 = M^2 - M; \quad (9)$$

$$P = P_1 + P_2 = 2M - 1 + \frac{3}{2}(M^2 - 2M + 1) = \frac{3}{2}M^2 - M + \frac{1}{2}. \quad (10)$$

According to formulas (5-10), graphs of the dependences of the number of additions S and multiplications P required to solve the Yul-Walker equations on the order of the autoregression model M were constructed.

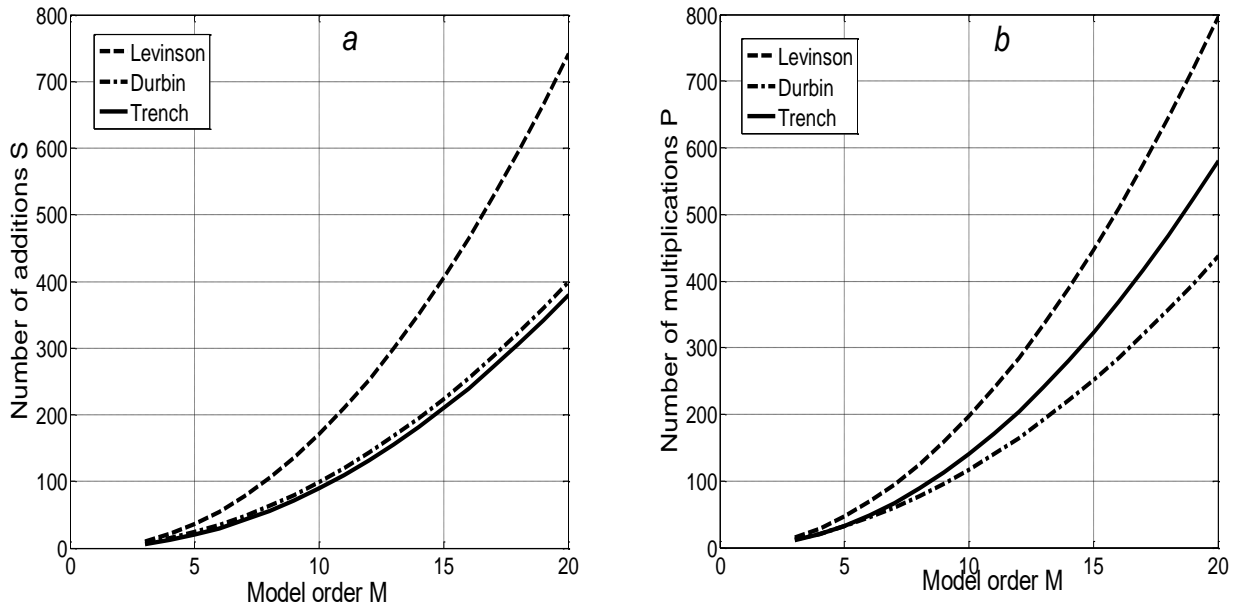


Fig 1. Dependence of the number of additions (a) and multiplications (b) on the order of the autoregression model

From the figures it is obvious that the number of additions is minimal for the Trench algorithm, the number of multiplications – for the Durbin algorithm. A more accurate comparison of the performance of algorithms can be performed taking into account the specific element base used.

Estimation of complexity of microcontroller realization of the solution of Yul-Walker equations

When creating portable spectrum analyzers for diagnostics of energy facilities and other technological

equipment, it is advisable to use hardware and software based on microcontrollers [12, 13]. For a reasonable choice of parameters of the microprocessor spectrum analyzer can be useful to calculate the approximate time of solving Yule-Walker equations on a microcontroller base, taking into account the previously obtained relations (5-10) and information about the time of arithmetic operations, given, for example, in [15].

Thus, the Mega 2560 microcontroller requires 12 microseconds for 64-bit integer addition and 60 microseconds for integer multiplication. For floating-point arithmetic operations, 13 and 15 microseconds,

respectively. As a result, the dependences of the time of solving the Yul-Walker equations on the order of the model were constructed (fig. 2). The obtained time

characteristics are, of course, evaluative, as they do not take into account the delay and branching commands, but can be the basis for a reasoned choice of technical means.

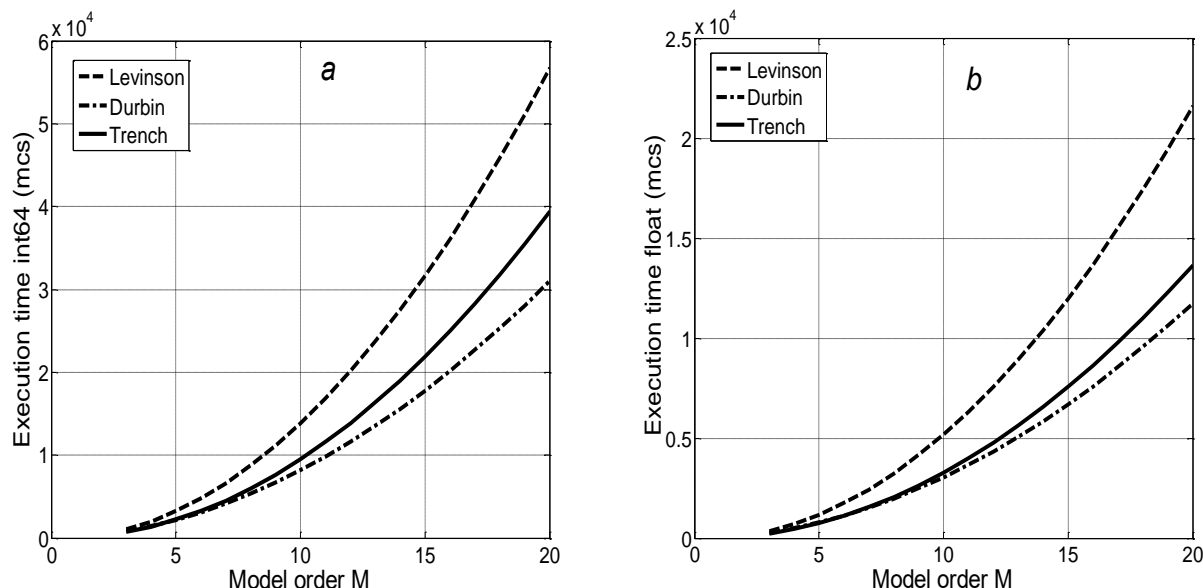


Fig. 2. Dependence of the time of solving the Yul-Walker equations on the model order (*a*-integer arithmetic, *b*-floating-point arithmetic)

It follows from the figures that the Durbin algorithm has the highest speed for fixed-point arithmetic. For floating-point arithmetic at small values of the model order ($M \leq 6$), the Trench algorithm is more efficient. Thus, for $M=5$ the calculation time according to the algorithm of the Trench algorithm is 755 μ s, according to the Durbin algorithm - 792 μ s. At values of $M > 6$ the performance of the Durbin algorithm is higher. Naturally, in the presence of data on the time of arithmetic operations, similar calculations can be performed on other platforms.

Conclusions

1. The spectral estimation method based on the Yul-Walker equations, which require the calculation of autoregression coefficients, combines sufficient resolution and noise immunity with acceptable implementation complexity.

2. Comparison of Levinson, Durbin and Trench algorithms for calculating autoregression coefficients showed that the Trench method requires a minimum number of additions, and the Durbin method – a minimum number of multiplications.

3. In the microcontroller implementation of spectrum analyzers should take into account the features of the arithmetic used by the controller. The Trench method is the fastest in the case of floating-point arithmetic and small-scale modeling. In other cases, Durbin's method is more effective.

4. In the future it would be actual to obtain estimates of the execution time of other components of spectral estimation algorithms – ACF calculation, solving Yul-Walker equations and Fourier transform calculation, which would allow a comprehensive assessment of software and hardware complexity of spectrum analyzers.

References

1. Shmelov, Y. M., Vladov, S. I., Kryshan, O. F., Gvozdik, S. D., Chyzhova, L. I. (2018), "Research of classification method of TV3-117 engine rating perations based on neural network technologies", ["Rozrobka metodu klasyfikatsii rezhyimiv roboty aviatsiinoho dvyhuna TV3-117 na osnovi neiromerezhevykh tekhnolohii"], *Innovative Technologies and Scientific Solutions for Industries*, No. 4, P. 93–102. DOI: <https://doi.org/10.30837/2522-9818.2018.6.093>
2. Raskin, L., Sukhomlyn, L., Ivanchikhin, Y., Korsun R. (2021), "Method of identification of object states according to the results of fuzzy measurements of controlled parameters", ["Metod identyfikatsii staniv obyektu za rezultatsy nechitkykh vymiriuvan kontroliuvanykh parametriv"], *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (18), P. 75–86. DOI: <https://doi.org/10.30837/ITSSI.2021.18.075>
3. Follum, J., Becejac, T., Etingov, P. A. (2021), "Robust Yule-Walker Method for Online Monitoring of Power System Electromechanical Modes of Oscillation", *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. DOI: <https://doi.org/10.1109/ISGT49243.2021.9372152>
4. TMForum (2021), "Application Framework (TAM)", available at: <https://www.tmforum.org/application-framework> (last accessed 15.12.2021).
5. TMForum (2021), "Information Framework (SID)", available at: <https://www.tmforum.org/information-framework-sid>. (last accessed 15.12.2021).
6. TMForum (2021), "Open Digital Architecture (ODA)", available at: <https://www.tmforum.org/oda/> (last accessed 15.12.2021).

7. Goel, S. (2020), "The journey from CSP to DSP. Informa PLC", available at: <https://telecoms.com/opinion/the-journey-from-csp-to-dsp/> (last accessed 15.12.2021).
8. Torosyan, E. K., Shekhovtsova, E. I., Akopyan, A. A. (2019), "Analysis of end-to-end methods for the basic architecture of telecom operator systems automation" ["Analiz metodov skvoznoy avtomatizatsii bazovoy arkhitektury sistem operatora svyazi"], *Petersburg economic journal*, No. 2., P. 172–177.
9. Newman, M. (2020), "FUTURE BSS. Say goodbye to software customization", available at: <https://inform.tmforum.org/research-reports/future-bss-say-goodbye-to-software-customization/> (last accessed 15.12.2021).
10. TMForum (2020), "Cloud native OSS/BSS", available at: <https://inform.tmforum.org/research-reports/cloud-native-oss-bss/> (last accessed 15.12.2021).
11. IIDA (2015), "A Guide to the Business Analysis Body of Knowledge (BABOK Guide)", IIBA, P. 512.
12. Wiegers, K. I., Bitti, J. (2019), "Software Requirements (Developer Best Practices) 3rd Edition" ["Razrabotka trebovaniy k programmnomu obespecheniyu. 3-e izdanie, dopolnennoe"], *BHV*, Spb., P. 736.
13. Mortensen, M. H., Offredo-Zreik, L. (2020), "Microservices-Based Cloud Native Modernization of OSS/BSS with Open Source. Informa PLC", available at: <https://telecoms.com/intelligence/microservices-based-cloud-native-modernization-of-oss-bss-with-open-source/> (last accessed 15.12.2021).
14. McElligott, T. (2019), "Future OSS. Toward an Open Digital Architecture. TMForum", available at: <https://inform.tmforum.org/research-reports/future-oss-towards-an-open-digital-architecture/> (last accessed 15.12.2021).
15. Smidovych, L.S., Kalmykov, A.V. (2010), "Perspective directions for the development of business processes automation technologies" ["Perspektivnye napravleniya razvitiya tekhnologiy avtomatizatsii biznes-protsessov"], *Radioelectronic and Computer Systems*, Vol. 4 (45), P. 75–81.
16. Turkington, I. (2021), "Open Digital Framework, Open Digital Architecture. ITU-T". *Virtual forum on The Role of Standards in Accelerating Digital Transformation for Cities and Communities*, P. 42.
17. Andersen, L. (2021), "Designing lasting simplification into a BSS contract. TMForum", available at: <https://inform.tmforum.org/digital-transformation-and-maturity/2021/08/designing-lasting-simplification-into-a-bss-contract/> (last accessed 15.12.2021).
18. Newman, M. (2021), "Next generation BSS: the Telefónica way. TM Forum", available from: <https://inform.tmforum.org/research-reports/next-generation-bss-operator-and-vendor-perspectives> (last accessed 15.12.2021).
19. Fedorovich, O.Ye., Babich, A.V. (2016), "Optimization of multilevel component architecture software system" ["Optimizatsiya mnogourovnevnoy komponentnoy arkhitektury programmnoy sistemy"], *Information Processing Systems*, 2016. Vol. 1 (75), P. 175–177.

Received 23.03.2022

Відомості про авторів / Сведения об авторах / About the Authors

Зуєв Андрій Олександрович – кандидат технічних наук, доцент, завідувач кафедри автоматики та управління в технічних системах, Національний технічний університет "Харківський політехнічний інститут", вул. Кирпичова, 2, Харків, Україна; e-mail: dakarton@gmail.com; ORCID: <http://orcid.org/0000-0001-8206-4304>

Зуєв Андрей Александрович – кандидат технических наук, доцент, заведующий кафедрой автоматики и управления в технических системах, Национальный технический университет "Харьковский политехнический институт", ул. Кирпичева, 2, Харьков, Украина.

Zuev Andrey – PhD, Associate Professor, Department of automation and control in technical systems, head of the department, National Technical University "Kharkiv Polytechnic Institute" Kyrpychova str., 2, Kharkov, Ukraine.

Івашко Андрій Володимирович – кандидат технічних наук, доцент кафедри автоматики та управління в технічних системах, Національний технічний університет "Харківський політехнічний інститут", вул. Кирпичова, 2, Харків, Україна; e-mail: ivashkoauts@gmail.com; ORCID: <https://orcid.org/0000-0002-4012-1697>

Ивашко Андрей Владимирович – кандидат технических наук, доцент кафедры автоматики и управления в технических системах, Национальный технический университет "Харьковский политехнический институт", ул. Кирпичева, 2, Харьков, Украина.

Ivashko Andrey – PhD, Associate Professor, Department of automation and control in technical systems, National Technical University "Kharkiv Polytechnic Institute", Kyrpychova str., 2, Kharkov, Ukraine.

Лунін Денис Олександрович – старший викладач кафедри автоматики та управління в технічних системах, Національний технічний університет "Харківський політехнічний інститут", вул. Кирпичова, 2, м. Харків, Україна, 61002; e-mail: luninden77@gmail.com; ORCID: <http://orcid.org/0000-0002-9418-0000>

Лунин Денис Александрович – старший преподаватель кафедры автоматики и управления в технических системах, Национальный технический университет «Харьковский политехнический институт», ул. Кирпичева, 2, Харьков, Украина.

Lunin Denis – Senior Lecturer Department of Automation and Control in Technical Systems National Technical University «Kharkiv Polytechnic Institute», Kyrpychova str., 2, Kharkiv, Ukraine.

ОЦІНКА ПРОГРАМНОЇ СКЛАДНОСТІ ОБЧИСЛЕННЯ КОЕФІЦІЄНТІВ АВТОРЕГРЕСІЇ ПРИ ЦИФРОВОМУ СПЕКТРАЛЬНОМУ АНАЛІЗІ

Предметом дослідження у статті є алгоритми швидкого обчислення коефіцієнтів авторегресії при цифровому спектральному аналізі та оцінки кількості арифметичних операцій, необхідних для їх реалізації. **Мета** статті – порівняльний аналіз

швидкодії різних алгоритмів обчислення коефіцієнтів авторегресії як складової частини алгоритмів спектрального аналізу, у тому числі аналіз складності їхньої мікроконтролерної реалізації. **Завдання**, що вирішуються: виділення методів спектрального аналізу, придатних для діагностики технологічного обладнання, аналіз методів обчислення коефіцієнтів авторегресії та виведення співвідношень для оцінювання програмної складності алгоритмів та розрахунок числових оцінок числа додавань та множень для деяких алгоритмів, адаптація розроблених методів та оцінок до мікроконтролерної реалізації аналізаторів спектру. Застосовувані **методи**: теорія алгоритмів, перетворення Фур'є, натуральні ряди, програмування мікроконтролерів. Отримані **результати**: показано, що методи спектрального оцінювання на основі рівнянь Юла-Уолкера, що вимагають обчислення коефіцієнтів авторегресії, поєднує достатню роздільну здатність і стійкість до перешкод з прийнятною складністю реалізації. Отримано оцінки числа додавань та множень для алгоритмів Левінсона, Дурбіна та Тренча, виконано їх порівняльний аналіз. Розраховані часи обчислення для мікроконтролерної арифметики з фіксованою та плаваючою точкою. **Висновки**: При побудові аналізаторів спектра для діагностики технологічного обладнання доцільно застосовувати метод Юла-Уолкера. Порівняння алгоритмів Левінсона, Дурбіна та Тренча для обчислення коефіцієнтів авторегресії показало, що метод Тренча потребує мінімальної кількості додавань, а метод Дурбіна – мінімальної кількості множень. При мікроконтролерній реалізації аналізаторів спектра слід враховувати особливості використовуваної контролером арифметики. Метод Тренча є найбільш швидкодіючим у разі застосування арифметики з плаваючою точкою та невеликого порядку моделі. В інших випадках ефективніший метод Дурбіна.

Ключові слова: спектральний аналіз; авторегресія; алгоритми Левінсона; Дурбіна; Тренча; обчислювальна складність; комп'ютерна арифметика; мікроконтролери.

ОЦЕНКА ПРОГРАММНОЙ СЛОЖНОСТИ ВЫЧИСЛЕНИЯ КОЭФФИЦИЕНТОВ АВТОРЕГРЕССИИ ПРИ ЦИФРОВОМ СПЕКТРАЛЬНОМ АНАЛИЗЕ

Предметом исследования в статье являются алгоритмы быстрого вычисления коэффициентов авторегрессии при цифровом спектральном анализе и оценки числа арифметических операций, необходимых для их выполнения. **Цель** статьи – сравнительный анализ быстродействия различных алгоритмов вычисления коэффициентов авторегрессии как составной части алгоритмов спектрального анализа, в том числе анализ сложности их микроконтроллерной реализации. Решаемые **задачи**: выделение методов спектрального анализа, пригодных для диагностики технологического оборудования, вывод соотношений для оценивания программной сложности алгоритмов и расчет числовых оценок вычислительной сложности, адаптация разработанных методов и оценок к микроконтроллерной реализации анализаторов спектра. Применяемые **методы**: теория алгоритмов, преобразование Фурье, линейная алгебра, натуральные ряды, программирование микроконтроллеров. Полученные **результаты**: показано, что метод спектрального оценивания на основе уравнений Юла-Уолкера, требующих вычисления коэффициентов авторегрессии, сочетает достаточную разрешающую способность и помехоустойчивость с приемлемой сложностью реализации. Получены оценки числа сложений и умножений для алгоритмов Левинсона, Дурбина и Тренча и выполнен их сравнительный анализ. Рассчитаны времена вычисления для микроконтроллерной арифметики с фиксированной и плавающей запятой. **Выводы**: сравнение алгоритмов Левинсона, Дурбина и Тренча для вычисления коэффициентов авторегрессии показало, что метод Тренча требует минимального количества сложений, а метод Дурбина – минимального количества умножений. При микроконтроллерной реализации анализаторов спектра следует учитывать особенности используемой контроллером арифметики. Метод Тренча является наиболее быстродействующим в случае применения арифметики с плавающей точкой и небольшого порядка модели. В остальных случаях более эффективен метод Дурбина.

Ключевые слова: спектральный анализ; авторегрессия; алгоритмы Левинсона; Дурбина; Тренча; вычислительная сложность; компьютерная арифметика; микроконтроллеры.

Бібліографічні описи / Bibliographic descriptions

Зуєв А. О., Івашко А. В., Лунін Д. О. Оцінка програмної складності обчислення коефіцієнтів авторегресії при цифровому спектральному аналізі. *Сучасний стан наукових досліджень та технологій в промисловості*. 2022. № 1 (19). С. 85–91. DOI: <https://doi.org/10.30837/ITSSI.2022.19.085>

Zuev, A., Ivashko, A., Lunin, D. (2022), "Estimation of software complexity of calculation of autoregression coefficients at digital spectral analysis", *Innovative Technologies and Scientific Solutions for Industries*, No.1 (19), P. 85–91. DOI: <https://doi.org/10.30837/ITSSI.2022.19.085>

V. KNYAZEV, B. LAZURENKO, A. SERKOV

METHODS AND TOOLS FOR ASSESSING THE LEVEL OF NOISE IMMUNITY OF WIRELESS COMMUNICATION CHANNELS

The **subject** of this article is the process of increasing the noise immunity of wireless channels in the construction or operation of modern digital communication systems (DCS), when the transmission of digital information significantly complicates the presence of interference and distortion in the communication channel. The **aim** is analysis and development of methods for ensuring stable and reliable operation of ultra-wideband wireless communication system in the conditions of interference and distortion of information in wireless communication channels. The **task** is creation and implementation of practical recommendations for improving the level of secrecy and quality of information circulating in the communication channel; development of criteria for quality, information efficiency and noise immunity of DCS wireless communication channels in the conditions of interference of natural and artificial origin. The **methods** used: methods of analytical modeling and the theory of potential noise immunity. The following **results** were obtained. The feasibility of assessing the DCS efficiency by two indicators - energy and frequency efficiency, which are the components that make up the overall integral indicator of informational efficiency was shown. It was proposed to evaluate the DCS efficiency by comparing the indicators of informational efficiency of the inverse and reference systems, using the Shannon limit as an ideal ratio of energy and frequency efficiency. On the basis of the theory of potential noise immunity it was shown that it is possible to operate wireless communication channels in conditions when the level of information signal and noise have the same value. **Conclusions.** It is shown that the use of the technology of supersmooth signals allows to carry out wireless redundant transmission of information with low impromptu capacity. Moreover, the efficiency of digital communication systems significantly increases due to the creation of an ensemble of complex signals through one-time encoding and modulation, which expands its information base, ensuring coverage and energy efficiency in a double symmetric channel. This enables the fullest possible utilization of the channel's Shannon capacity with high integrity of information transfer without a significant increase in the signal-to-noise ratio at the receiver's input. An integral indicator of informational efficiency of the wireless DCS communication channel is provided. It enables real-time integral assessment of link quality and optimal routing in a mobile peer-to-peer network – MANET.

Keywords: Information effectiveness; noise protection; noise immunity; concealment; communication channel.

Introduction

The amount of information is increasing every year, the range of communication is increasing, and the requirements for information quality are becoming higher. Wireless transmission is severely hampered by the presence of obstacles and spoils in the communication channel [1 - 3]. Wireless networks of the next generation are becoming extremely brittle and multifaceted. The use of unmanned aerial vehicle as air base stations of mobile communication makes it possible to support data transmission with high speed for users, in the absence of typical cellular infrastructure [4, 5]. However, the use of these mobile devices, which move in a dynamic space, significantly complicates the problems of mobility, since the relocation to a new location leads to the possible disconnection of the current users. The lack of infrastructure is the main characteristic of mobile peer-to-peer networks - MANET. Therefore, mobile devices that are part of a peer-to-peer network must be able to communicate with one another, acting as routers. They move in space unnecessarily, changing the topology of the network, which requires constant adaptation and reconfiguration of routes, so that the devices will be able to interact with each other. At the same time, a reliable stable connection between devices should be ensured in real time to maintain the appropriate level of service quality and interoperability, reducing delays in data delivery between mobile devices. In view of these requirements, there are methods for controlling the parameters of information and telecommunication networks of critical use [6], re-routing of base stations in mobile communication networks [7], re-routing of data traffic in infocommunication networks [8] and throughput

capacity in software-reconfigured networks [9] should be improved to increase reliability of communication and handlers between devices and equipment. Due to the high mobility the information delay increases, wireless communication deteriorates, which causes difficulties in maintaining communication channels of higher quality. Thus, to determine the current optimal topology of the network and quality link channels, an integral indicator of informational efficiency should be developed, which will reduce the time of making an optimal decision. Under these conditions, the problem of increasing noise immunity and efficiency of communication channels becomes one of the most important problems of current wireless communication theory and technology in the construction and operation of modern DCS.

Quality indicators of wireless mobile communication channels

Under conditions of constant growth of demand in communication services and severe limitation of frequency resource there is a need for more efficient use of it. In this case the combination of transmitter, receiver and communication lines creates a communication channel. As the *noise immunity of communication channels* characterizes the ability of the system to ensure the transmission of messages with a given quality under conditions of natural and piecemeal interferences, it also characterizes its ability to resist the means of radio-electronic suppression. Therefore, the noise immunity of communication channels should be considered as the summation of two components – noise immunity and concealment. We understand the *channel concealment* as its energy concealment, including the ability to store the

fact of information transmission, structural concealment of the signal and the way of its encoding. Since these properties and factors are of an accidental nature, noise immunity and concealment must be defined by the same criteria. When a channel does not have the property of probability or noise immunity, it does not have the property of protection. Thus, noise immunity of a channel is defined by two channel properties – concealment and noise immunity.

The quality of operation of digital communication systems is characterized by a set of indicators, of which the main ones are the accuracy of renewal of messages, noise immunity, speed, throughput capacity, range of operation, the level of electromagnetic coupling and others. The sum of system quality indicators is recorded in the form of a vector:

$$Q = (Q_1, Q_2, \dots, Q_n).$$

The comparison of systems according to these indicators allows to identify the highest quality system, which corresponds to the highest / lowest value of a certain function of partial quality indicators.

$$Q = \varphi(Q_1, Q_2, \dots, Q_n).$$

Thus, the value Q is the efficiency or a generalized indicator of the system's performance, and $\varphi(Q_1, Q_2, \dots, Q_n)$ is the objective function of the system. However, such a large set of indicators does not allow us to selectively choose the most efficient variant. Therefore, we should develop a universal integral indicator of binary communication channel quality assessment for its use both at the stage of development and at the stage of DCS operation.

The main structural element of DCS is the communication channel. The maximum efficiency of the channel occurs when the information signal meets the requirements of the communication channel in the best way. Based on the analysis of communication systems [10 – 11], the main indicator of the quality is acceptable to take the authenticity of transmitted information. For this purpose the error coefficient P_e is often used – the probability of error during the transmission of a unit of information:

$$P_e = Q\left(\sqrt{\frac{E}{N_0}}\right),$$

where: $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$ is a Gaussian error probability integral.

Another indicator of quality is the signal-to-noise ratio $q_0 = \frac{E_s}{N_0}$, which relates the energy of the signal (bit)

E_s to the spectral intensity of the noise N_0 .

When assessing the quality of the transmitted information, the following parameters are often used: P_s – the probability of error and the signal/noise ratio q_0 , which is due to their unambiguous mutual dependence for a particular DCS. The strength of the signal and the bandwidth (signal base) are the main resources that are expended on information transfer. Therefore, it is reasonable to reduce the level of information signal at the input of the receiver to the noise level in order to ensure a smooth and fault-free operation. This determines the appropriate level of reception and noise immunity of wireless communication channels (fig. 1).

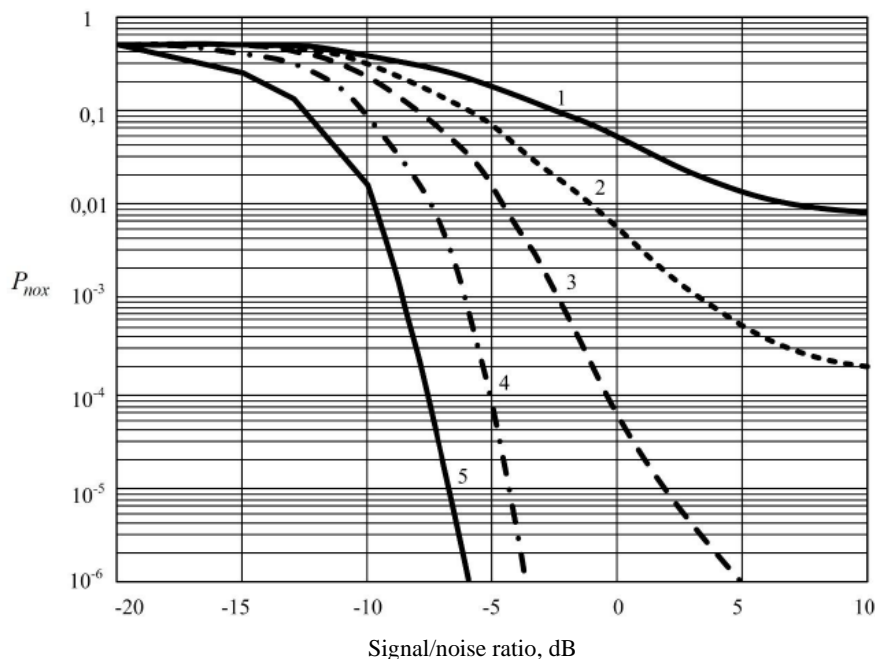


Fig. 1. The dependence of the probability of the error rate P on the signal-to-noise ratio at different signal bases: (1-B=50; 2-B=100; 3-B=200; 4-B=500; 5-B=1000).

Characteristics analysis has shown that at a high signal base $B > 300$, when the intensity of the received signals is lower than the level of impairment, information transfer is carried out with a breakeven of less than 10-6, which characterizes its high probability of reliability. For digital double-beam signals the reliability of the received information is increased by 10 - 20 dB at low (-3, ..., -5 dB) signal/noise ratios by increasing the basis of the information signal [12, 13]. The use of orthogonal functions as channel signals is due to the fact that the separation of these signals is performed without degradation of the signal/noise ratio.

Extension of information base is most effectively realized through creation of complex signal ensemble by means of its one-time coding and modulation [14], which allows the most complete use of the channel's bandwidth with high authenticity of information transmission without a significant increase in signal/noise at the receiver's input.

When a complex signal is used, when a bit of information is transmitted during the period T_s by a binary numeric sequence function, its base is defined by the following relationship:

$$B_s = T_s / \tau_s$$

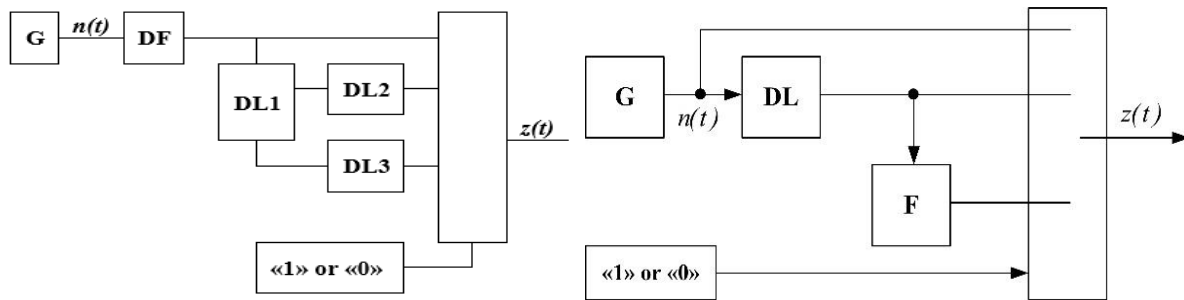


Fig. 2. Scheme of formation an ensemble of complex LNS information signal

In the middle of the bit interval the commutator is switched to one of the two possible positions, depending on the flow of two-way bits "one" or "zero" from the information source. Thus, the delay line DL1 ensures delay of the signal $x(t)$ by half of the bit interval $T_s/2$, and delay lines DL2 and DL3 are backed indirectly to form a stream of double bits "one" or "zero". As a result, an ensemble of complex signals is formed, where the information signal, separated in time from the reference signal, and on one interval T_s has the following form.

$$y(t) = \begin{cases} x(t), & 0 \leq t \leq \frac{T}{2}; \\ x\left(t - \frac{T}{2} - T_0\right), & \frac{T}{2} \leq t \leq T; \\ x\left(t - \frac{T}{2} - T_1\right), & \frac{T}{2} \leq t \leq T. \end{cases}$$

Thus, the use of LNS communication technology in the timing of the synchronous pulse and information signal allows to ensure that the requirements of noise immunity, security and safety of wireless mobile

where τ_s is a duration of the elementary symbol (chip) of the numerical sequence.

Thus, increasing the signal base while preserving the speed of information transmission is possible by reducing the duration of the chip τ_s .

At the same time, the wide frequency band and ultra-short duration of the chips make increased demands on the accuracy of synchronization. Information signals and synchronization signals are at the same energy level, and the spectral density of all channels is at the noise level, so the system is significantly degraded by the values of the probability of bit error. Thus, the simultaneous coding, synchronization and modulation of the information bit is carried out, creating a complex signal-code structure. As a result, an ensemble of the complex signal is formed, where the information signal is isolated in time from the reference one bit interval T_s . The transmitter circuit (fig. 2) uses a tripping commutator, which switches the transmitter output directly to the LNS signal generator (G) during the first half of the bit interval. Thus, during the period of time $T_s/2$ the reference LNS signal is formed.

communication channels at all stages of their development, production and operation are guaranteed.

Due to the creation of complex signal-code structures of information signal by a one-time use of processes of coding and modulation receives a signal that occupies a greater range of frequencies, and has significantly less of its intensity in comparison with the signal that is realized by their successive processes. This allows for higher efficiency and energy gain. Spectrum broadening results in the signal being distributed more evenly and less narrowly in the specified range of the spectrum. Thus, not only increases the noise immunity of information, but also reduces the likelihood of its overexposure due to the absence of a synchronized copy of the broadening signal. Considering that the problem of organization of high-quality mobile communication and protection of information in wireless radio networks is to reduce the level of backward electromagnetic environment, the technology of Ultra-Wide band signals is the most appropriate for its practical use. However, the evaluation of the efficiency of these or other technological solutions when implementing them in DCS requires the use of certain evaluation criteria.

Method for increasing the energy efficiency of digital communication systems

Concealment of communication channels is characterized by its energetic concealment, the components of which are the processes of forming, coding and modulation of information signal. However, the universal indicator is the base of the information signal, the value of which determines the level of availability of the communication channel.

The main way to increase energy earliness of the communication channel of the digital wireless system is to decrease the signal/noise ratio at the output of the correlation receiver. The dependence of the variability of the hysteresis P on the signal-to-noise ratio at different signal bases (fig. 1) leads to the tendency to the maximum possible reduction of the signal power intensity to the necessity of using the signal at the maximum base. The dependence of the quality of information transmission on the width of the signal spectrum leads to the fact that DCS with complex signals considerably increases the s/n ratio at high demands to the quality of information transmission.

According to V.A. Kotelnikov's potential noise immunity theory [15], the limit for the reduction of the information signal level for all classes of receiver systems without exception is the ratio of the subsurface energy of the signal E to the spectral power density of noise N_0 , which is the value:

$$Q = 2E/N_0 = 2q_0 B,$$

where: $q_0 = \frac{E/T}{N_0 W}$ is the ratio of the average signal power

$P_{s0} = E/T$ to the noise power $P_{N0} = N_0 W$ at the receiver input, and $B = WT$ is the signal basis.

In the systems of radio-electronic antidetection when radio contamination of the communication channel is carried out, first of all, the presence of corrupting signal is detected. However, the intelligence receiver practically does not detect the useful signal $q_0 \leq 1$ when it is used. This is due to the fact that its square detector further reduces the useful signal that will be identified as interference. The resulting signal/noise value at the output of the receiver detector for a uniform spectral signal strength is determined by the following ratio:

$$\left(\frac{P_s}{P_n} \right)_{out} \approx 55q_0^2,$$

where: P_s is a signal power, P_n is a noise power.

At the same time, the following relationship will take place for the values $q_0 \geq 1$:

$$\left(\frac{P_s}{P_n} \right)_{out} \approx 4,6q_0.$$

Thus, at a weak signal in the intelligence receiver there is a suppression of the useful signal by interference, which significantly increases its energy concealment.

Method and criteria for evaluating the noise immunity of the digital communication system

Noise immunity of the discrete communication channel depends on the structure and type of signal and interferences, their intensity, the structure of the channel receiver, type of antennas, etc. However, if the optimal signal reception is performed against the background of the own noise of the receiver, the noise immunity of signal reception of any shape will be the same and will be determined only by the signal energy and spectral noise power, which is brought to the input of the receiver. DCS technical effect determines the amount and quality of transmitted information for a certain period of time or per unit of time, i.e. transmission rate R (bit/s) and the probability of error P in the transmission of discrete messages.

Indicators of energy $\beta = \frac{R_u N_0}{P_s}$ and frequency

$\gamma = \frac{R_u}{F_s}$ efficiency, determine the consumption of these resources. In these relations P_s is a signal power, F_s - signal bandwidth, R_u - transmission speed.

Combining these indicators, we get the integral indicator of informational efficiency:

$$\eta = \frac{R_u}{C},$$

where C is a communication channel capacity.

It is the most general indicator, which determines the technical efficiency of the system. Graphical dependence for the channel efficiency is presented in coordinates $\beta \gamma$ in fig. 2. It is a limit and mirrors the best exchange between β and γ . In coordinates β and γ , each variant of the real system will be matched by the corresponding point on the plane, which is located below the Shannon limit dependence [16]. When comparing it with the ideal system (Shannon's limit) the winnings will be negative. It determines how close the inverse system is to the ideal system.

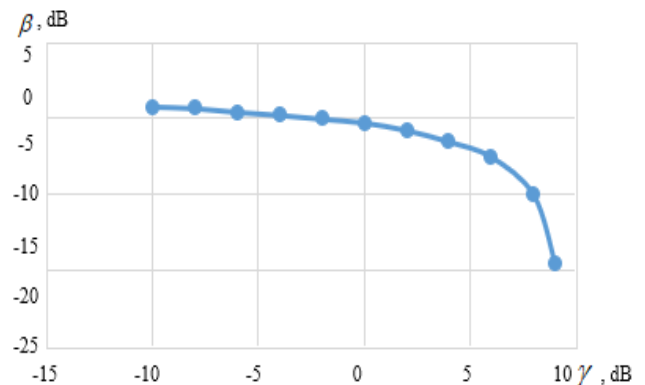


Fig. 3. Shannon's limit

The coefficients β and γ can be calculated for most real systems in order to compare them. Varying these interrelated characteristics, there is a possibility to estimate the degree of influence of characteristics of real channels and non-ideality of parameters of transmission and reception devices and to determine their optimum correlation

$$\beta = \frac{\gamma}{2^{\gamma} - 1}.$$

Under the condition $\eta = 1$ we obtain the boundary dependence between β and γ (Shannon limit), which means that any radio channel will have poorer characteristics of energy and frequency efficiency.

Channel throughput with additive Gaussian noise (in bps).

$$C = \log \left(\frac{P_s}{P_n} + 1 \right). \quad (1)$$

In calculation per character (bits/countdown):

$$C_n = 0,5 \log \left(\frac{P_s}{P_n} + 1 \right).$$

The channel capacity utilization coefficient (informational efficiency) is entered as a summary indicator of DCS technical efficiency.

$$\eta = \frac{R_u}{C} = 1 - \chi, \quad (2)$$

where: $\chi = \frac{(C - R_u)}{C} = 1 - \eta$ is a redundancy of the channel.

According to (1) and (2)

$$\eta = \gamma \log \left(\frac{\gamma}{\beta} + 1 \right).$$

We take this function as a target and determine its maximum value. The value η can be close to one, and the error can be arbitrarily small. In real systems, the error always has a finite value and $\eta < 1$. In these cases, β and γ are determined separately and the dependences $\beta = f(\gamma)$ are constructed for $p = const$. It should be noted that in digital binary communication systems, the coefficient γ can vary from 0 to $\log 2 = 0.3$ dB, and the coefficient β from 0 to $\frac{1}{\ln 2} = 1.44$ dB.

Digital communication systems are divided into two groups. These are systems with high β but low γ efficiency. Digital communication systems belong to this group, which must ensure the best possible use of the signal power at a given accuracy of transmission. In particular, this concerns space and satellite communication systems. The other group includes systems with high γ , but low β – efficiency, which is typical for the systems of conducting communication, when it is necessary to obtain the best possible use of the channel frequency smoothness at a given transmission variability. Sometimes, a compromise solution should be

adopted, which allows to obtain simultaneously the values of β and γ of Diagram (fig. 3) to solve this problem. For wireless communication channels the technical effect is conveniently determined by the energy gain:

$$\Delta\beta = \beta / \beta_s \text{ при } \gamma = \gamma_{add.},$$

where β is an energy efficiency of the selected system;

β_s is an energy efficiency of the basic (reference) system. Thus, the obtained β/γ -diagrams allow turning DCS, which meet the specified requirements, or to optimize for the coefficients of β and γ .

Analysis

Analysis of the problem of ensuring noise immunity and reliability of wireless communication channels has shown that the main method of its solution is to reduce the level of electromagnetic field impromptu. In the case when the speed of information transmission is set ($R_u = const$), it is possible to operate at a low signal/noise ratio, which is a common problem in a faulty communication system, only by decreasing the frequency and energy efficiency. In this case, according to Shannon's theorem, due to the appropriate coding the quality of transmission of messages can be made as good as desired. On the basis of the theory of potential noise immunity Kotelnikov V.A. shows the possibility of the established and warrantless operation of wireless digital communication system in conditions where the level of information signal and noise have the same value. The possibility of separating the transmitted information signal from the noise by correlation of the received and reference signals is shown.

Conclusions

To increase the energy capacity and noise immunity of the communication channel, it is necessary to use the transmission with the lowest possible quality indicator q_0 . In this case, it is advisable to use highly directional antennas with the smallest possible level of lateral pendulums, as well as to use a noise signal with the highest value of the signal base $B \gg 1$, using complex signals.

It is shown that the use of the technology of supersmooth signals allows to carry out wireless redundant transmission of information with low impromptu capacity. Moreover, the efficiency of digital communication systems significantly increases due to the creation of an ensemble of complex signals through one-time encoding and modulation, which expands its information base, ensuring concealment and energy efficiency in a dual symmetric channel. Diffusion in the timing of the synchropulse and information signal within the limits of the duration of the information beat allows you to securely ensure the requirements of noise immunity, prihozanovosti and safety of wireless communication channels at all stages of their design, manufacture and operation at the expense of transmission of information signals with low impromptu power. Using a large base signal makes it possible to ensure smooth and seamless operation of the digital communication system

under conditions where the level of information signal is at or below the noise level. The Integral Indicator of Informational Efficiency of Wireless DCS

Communication Channel was provided. It enables real-time integral assessment of link quality and optimal routing in a mobile peer-to-peer network - MANET.

References

- Uriadnikov, Yu.F. (1991), *Theory of interference-proof radio control and communication channels [Teoriia pomechozashchishchennikh radiokanalov upravleniia i svyazi]*, MO SSSR, Moscow, 224 p.
- Gorbenko, I., Zamula, O., & Lik, Kh. Ch. (2020), "Comprehensive solution to the problem of electromagnetic compatibility of modern information and communication systems", *Radiotekhnika*, 3 (202), P. 106–115. DOI: <https://doi.org/10.30837/rt.2020.3.202.11>
- Moskalets, M., Loshakov, V., Abdenour, D., Ageyev, D., Martynchuk, O., Seliivanov, K. (2019), "Methods for Solving EMC Problems by Means of Antenna Technology in Tropospheric Communication", *IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*, P. 571–574. DOI: <https://doi.org/10.1109/PICST47496.2019.9061532>
- Ayass, T., Coqueiro, T., Carvalho, T., Jaiton, J., Araújo, J., Francês, R. (2022), "Unmanned aerial vehicle with handover management fuzzy system for 5G networks: challenges and perspectives", *Intell Robot*, 2 (1), P.20–36. DOI: <http://dx.doi.org/10.20517/ir.2021.07>
- Adzhemov, S.S., Sokolova, M.V., Uriadnikov, Yu.F., Shtirkin, V.V. (2006), "Ultra-broadband is the result of the development of broadband access technologies" ["Sverkhshirokopolosnaia svyaz – rezultat razvitiia tehnologii shirokopolosnogo dostupa"], *Elektrosvyaz*, No. 2, P. 18–23.
- Kosenko, V. (2017), "Principles and Structure of the Methodology of Risk Adaptive Management of Parameters of Information and Telecommunication Networks of Critical Application Systems", *Innovative Technologies & Scientific Solutions for Industries*, No. 1 (1), P. 46–52. DOI: <https://doi.org/10.30837/2522-9818.2017.1.046>
- Ruban, I., Kuchuk, H., Kovalenko, A. (2017), "Redistribution of Base Stations Load in Mobile Communication Networks", *Innovative Technologies & Scientific Solutions for Industries*, No. 1(1), P. 75–81. DOI: <https://doi.org/10.30837/2522-9818.2017.1.075>
- Lemeshko, O., Yevdokymenko, M., Yeremenko, O. (2019), "Model of Data Traffic QoS Fast Rerouting in Infocommunication Networks", *Innovative Technologies & Scientific Solutions for Industries*, No. №3 (9), P. 127–134. DOI: <https://doi.org/10.30837/2522-9818.2019.9.127>
- Yeremenko, O., Yevdokymenko, M., Sleiman, B. (2020), "Advanced Performance Based Fast Rerouting Model with Path Protection and its Bandwidth in Software Defined Network", *Innovative Technologies & Scientific Solutions for Industries*, No. 1 (11), P. 163–171. DOI: <https://doi.org/10.30837/2522-9818.2020.11.163>
- Knyazev, V., Kravchenko, V., Lazurenko, B., Serkov, O., Trubchaninova, K., Panchenko, N. (2022), "Development of methods and models to improve the noise immunity of wireless communication channels", *Eastern – European Journal of Enterprise Technologies*, No. 1/5 (115), P. 35–42. DOI: <https://doi.org/10.15587/1729-4061.2022.253458>
- Lazurenko, B., Serkov, O., Trubchaninova, K. (2020), "Security improvement techniques for mobile applications of Industrial Internet of Things", *IJCSNS International Journal of Computer Science and Network Security*, Vol. 20, No. 5, P. 145–149.
- Serkov, A., Trubchaninova, K., Mazzites, M. (2019), "Method of wireless transmission of digital information on the basis of ultra-wide signals", *Advanced Information Systems*, Vol. 3, No. 4, P. 33–38. DOI: <https://doi.org/10.20998/2522-9052.2019.4.04>
- Lazurenko, B.O., Pevnev, V.Y., Serkov, O.A., Tkachenko, V.A., Kharchenko, V.S. (2021), *The method of transmitting information over wide-space pulse signals: a patent for wine [Sposib peredachi informatsii nadshyrokosmuhovymy impulsnymy sygnalamy]*, Patent No. UA 123519 Ukraina, MPK H04B 1/02 (2006.01), MPK H04B 1/69 (2011.01) MPK H04B 7/00, opubl. 14.04.2021, Biul. No. 15.
- Morelos-Saragosa, P. (2005), *The art of error-correcting coding. Methods, algorithms, application [Iskusstvo pomechoustoichivogo kodirovaniia. Metodu, algoritmu, primeneniia]*, Technosfera, Moscow, 320 p.
- Kotelnikov, V.A. (1956), *Theory of Potential Interference Immunity [Teoriya potencial'noj pomekhoustojchivosti]*, Gosenergoizdat, Moscow, 152 p.
- Shannon, C.E. (1963), *Works on information theory in cybernetics [Raboti po teorii informatsii v kibernetike]*, Inostrannaia Literatura, Moscow, 829 p.

Received 23.03.2022

Відомості про авторів / Сведения об авторах / About the Authors

Князев Володимир Володимирович – науково-дослідний та проектно-конструкторський інститут "Молнія", канд. техн. наук, ст. наук. співробітник, керівник НДВ-2 "Електромагнітна сумісність та безпека", Харків, Україна; e-mail: knyaz2@i.ua; ORCID ID: <https://orcid.org/0000-0002-7119-7790>.

Князев Владимир Владимирович – научно-исследовательский и проектно-конструкторский институт "Молния", канд. техн. наук, ст. наук. сотрудник, руководитель НДВ-2 "Электромагнитная совместимость и безопасность", Харьков, Украина.

Knyazev Vladimir Vladimirovich – Reserch and Desing Institute "Molniya", PhD, Senior Resercher, Head of RD-2 "Electromagnetic Compatibility and Safety" Kharkiv, Ukraine.

Лазуренко Богдан Олександрович – Національний технічний університет "Харківський політехнічний інститут", Харків, Україна, аспірант програми "Телекомунікації та радіотехніка"; e-mail: torroloco789@gmail.com; ORCID ID: <https://orcid.org/0000-0002-1914-7091>.

Лазуренко Богдан Александрович – Национальный технический университет "Харьковский политехнический институт", аспирант программы "Телекоммуникации и радиотехника", Харьков, Украина.

Lazurenko Bogdan – National Technical University "Kharkiv Polytechnic Institute", Postgraduate of Program "Telecommunication and Radio Engineering", Kharkiv, Ukraine.

Серков Олександр Анатолійович – Національний технічний університет "Харківський політехнічний інститут", доктор технічних наук, професор, професор кафедри систем інформації імені В.О. Кравця, Харків, Україна; e-mail: aleksandr.serkov@hotmail.com; ORCID ID: <https://orcid.org/0000-0002-6446-5523>.

Серков Александр Анатольевич – Национальный технический университет "Харьковский политехнический институт", доктор технических наук, профессор, профессор кафедры систем информации имени В.А. Портного, Харьков, Украина.

Serkov Aleksandr – National Technical University "Kharkiv Polytechnic Institute", Doctor of Sciences (Engineering), Professor, Professor of Information Systems Department, Kharkiv, Ukraine.

МЕТОДИ І ЗАСОБИ ОЦІНКИ РІВНЯ ЗАВАДОСТІЙКОСТІ БЕЗПРОВІДНИХ КАНАЛІВ ЗВ'ЯЗКУ

Предметом розгляду статті є процеси підвищення завадостійкості безпроводних каналів при побудові чи експлуатації сучасних цифрових систем зв'язку (ЦСЗ), коли передачу цифрової інформації суттєво ускладнює наявність завад та спотворень в каналі зв'язку. **Мета** – аналіз та розробка методів забезпечення усталеної та надійної роботи надширокополосної системи безпроводного зв'язку в умовах дії завад та спотворень інформації в безпроводних каналах зв'язку. **Задача** – створення та впровадження практичних рекомендацій щодо покращення рівня прихованості і якості інформації, яка циркулює в каналі зв'язку; розробка критеріїв якості, інформаційної ефективності та завадостійкості безпроводних каналів зв'язку ЦСЗ в умовах дії завад природного та штучного походження. Використані **методи**: методи аналітичного моделювання та теорії потенційної завадостійкості. Отримані наступні **результати**. Показана доцільність проведення оцінки ефективності ЦСЗ за двома показниками – енергетичної та частотної ефективності, які є складовими та створюють загальний інтегральний показник інформаційної ефективності. Запропоновано здійснювати оцінку ефективності ЦСЗ шляхом порівняння показників інформаційної ефективності обраної та еталонної систем, використовуючи межу Шеннона як ідеальне співвідношення енергетичної та частотної ефективності. На ґрунті теорії потенційної завадостійкості показана можливість усталеної та беззавадової роботи безпроводних каналів зв'язку в умовах, коли рівень інформаційного сигналу та шуму мають одне значення. **Висновки**. Показано, що використання технології надширокополосних сигналів дозволяє здійснити безпроводну приховану передачу інформації з малою потужністю випромінювання. Причому ефективність цифрових систем зв'язку суттєво збільшується за рахунок створення ансамблю складного сигналу шляхом одночасного кодування і модуляції, що розширює його інформаційну базу, забезпечуючи прихованість та енергетичну ефективність у двійковому симетричному каналі. Це дозволяє найбільш повно використовувати шеннонівську пропускну здатність каналу з високою достовірністю передачі інформації без суттєвого підвищення відношення сигнал/шум на вході приймача. Запропоновано інтегральний показник інформаційної ефективності безпроводного каналу зв'язку ЦСЗ. Він дозволяє у режимі реального часу здійснювати інтегральну оцінку якості каналу зв'язку та виконувати оптимальну маршрутизацію в мобільній одноранговій мережі – MANET.

Ключові слова: інформаційна ефективність; завадозахищеність; завадостійкість; прихованість; канал зв'язку.

МЕТОДЫ И СРЕДСТВА ОЦЕНКИ УРОВНЯ ПОМЕХОУСТОЙЧИВОСТИ БЕСПРОВОДНЫХ КАНАЛОВ СВЯЗИ

Предметом рассмотрения статьи являются процессы повышения помехоустойчивости беспроводных каналов при построении или эксплуатации современных цифровых систем связи (ЦСЗ), когда передачу цифровой информации существенно усложняет наличие помех и искажений в канале связи. **Цель** – анализ и разработка методов обеспечения устойчивой и надежной работы сверхширокополосной системы беспроводной связи в условиях помех и искажений информации в беспроводных каналах связи. **Задача** – создание и внедрение практических рекомендаций по улучшению уровня скрытности и качества информации, циркулирующей в канале связи; разработка критериев качества, информационной эффективности и помехоустойчивости беспроводных каналов связи ЦСЗ в условиях действия помех естественного и искусственного происхождения. Используемые **методы**: методы аналитического моделирования и теории потенциальной помехоустойчивости. Получены следующие **результаты**. Показана целесообразность проведения оценки эффективности ЦСЗ по двум показателям – энергетической и частотной эффективности, которые являются составляющими и создают общий интегральный показатель информационной эффективности. Предложено оценку эффективности ЦСЗ путем сравнения показателей информационной эффективности выбранной и эталонной систем, используя границу Шеннона как идеальное соотношение энергетической и частотной эффективности. На основе теории потенциальной помехоустойчивости показана возможность установившейся и беспомеховой работы беспроводных каналов связи в условиях, когда уровень информационного сигнала и шума имеют одно значение. **Выводы**. Показано, что использование технологии сверхширокополосных сигналов позволяет осуществить беспроводную скрытую передачу информации с малой мощностью излучения. Причем эффективность цифровых систем связи существенно увеличивается за счет создания ансамбля сложного сигнала путем одновременного кодирования и модуляции, что расширяет его информационную базу, обеспечивая скрытость и энергетическую эффективность в двоичном симметричном канале. Это позволяет наиболее полно использовать Шенноновскую пропускную способность канала с высокой достоверностью передачи информации без существенного повышения отношения сигнал/шум на входе приемника. Предложен интегральный показатель информационной эффективности беспроводного канала связи ЦОС. Он позволяет в режиме реального времени производить интегральную оценку качества канала связи и выполнять оптимальную маршрутизацию в мобильной одноранговой сети – MANET.

Ключевые слова: информационная эффективность; помехозащищенность; помехоустойчивость; скрытность; канал связи.

Бібліографічні описи / Bibliographic descriptions

Князев В. В., Лазуренко Б. О., Серков О. А. Методи і засоби оцінки рівня завадостійкості безпроводних каналів зв'язку. Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19). С. 92–98. DOI: <https://doi.org/10.30837/ITSSI.2022.19.092>

Knyazev, V., Lazurenko, B., Serkov, A. (2022), "Methods and tools for assessing the level of noise immunity of wireless communication channels", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (19), P. 92–98. DOI: <https://doi.org/10.30837/ITSSI.2022.19.092>

АЛФАВІТНИЙ ПОКАЖЧИК

Альохіна С. В.	65
Багаєв І. О.	76
Барковська О. Ю.	5
Батюк Т. М.	13
Висоцька В. А.	13
Давидовський Ю. К.	47
Єлізева А. В.	40
Зуєв А. О.	85
Івашко А. В.	85
Князєв В. В.	92
Косенко В. В.	55
Лазуренко Б. О.	92
Лунін Д. О.	85
Мінухін С. В.	27
Момот М. О.	40
Настенко О. С.	5
Невлюдов І. Ш.	65
Попов А. В.	40
Ромашов Ю. В.	65
Семенов С. Г.	55
Серков О. А.	92
Смідович Л. С.	47
Сороколат Н. А.	76
Фатєєва Л. Ю.	76
Хомич В. М.	5
Цао Вейлін	55
Черняк О. М.	76

ALPHABETICAL INDEX

Alyokhina Svitlana	65
Ihor Bahaiev	76
Barkovska Olesia	5
Taras Batiuk	13
Victoria Vysotska	13
Davydovskyi Yurii	47
Yelizyeva Alina	40
Zuev Andrey	85
Ivashko Andrey	85
Knyazev Vladimir	92
Kosenko Viktor	55
Lazurenko Bogdan	92
Lunin Denis	85
Minukhin Serhii	27
Momot Myroslav	40
Nastenko Oleksandr	5
Nevliudov Igor	65
Popov Andrei	40
Romashov Yurii	65
Semenov Serhii	55
Serkov Aleksandr	92
Smidovych Leonid	47
Sorocolat Nataliia	76
Fatieieva Lina	76
Khomych Viktor	5
Cao Weiling	55
Cherniak Olena	76

НАУКОВЕ ВИДАННЯ

**СУЧАСНИЙ СТАН НАУКОВИХ ДОСЛІДЖЕНЬ ТА
ТЕХНОЛОГІЙ В ПРОМИСЛОВOSTІ**

Щоквартальний науковий журнал

№ 1 (19), 2022

Відповідальний секретар журналу *І. Г. Перова*
Відповідальний за випуск *А. А. Коваленко*
Комп'ютерна верстка *А. І. Коломієць*

АДРЕСА РЕДАКЦІЇ:

Україна, 61166, м. Харків, проспект Науки, 14

Тел.: +38 (057) 704-10-51

Веб-сайт: <http://itssi-journal.com>

E-mail: journal.itssi@gmail.com

Формат 60×84/8. Умов. друк. арк. 11,25. Тираж 150 прим.

Віддруковано з готових оригінал-макетів в типографії ФОП Андреев К.В.
Єдиний державний реєстр юридичних осіб та фізичних осіб-підприємців.
Запис №24800170000045020 від 30.05.2003.

61166, Харків, вул. Серпова, 4, тел. 063-993-62-73
e-mail: ep.zakaz@gmail.com