



MINISTRY
OF EDUCATION AND SCIENCE
OF UKRAINE



ACADEMY

OF TECHNICAL SCIENCES OF UKRAINE

CONNECTIVE
technologies

DEPARTMENT OF

INFORMATION

TECHNOLOGY

2022 International Conference on Innovative Solutions in Software Engineering (ICISSE)

Conference Proceedings

November 29-30, 2022
Ivano-Frankivsk, Ukraine

*Issued by the decision of the Academic Board of the Faculty of Mathematics and Computer Science
(protocol #15/2022)*

This book and its papers are published under the Creative Commons License Attribution 4.0 International
(CC BY 4.0)

Editors

Mykola Kuz  and **Mykola Kozlenko** 

Department of Information Technology
Faculty of Mathematics and Computer Science
Vasyl Stefanyk Precarpathian National University
57 Shevchenka Street
Administration Building, Office #319
Ivano-Frankivsk, 76018, Ukraine
Tel: +380 (342) 59-60-58
E-mail: icisse@pnu.edu.ua

Papers are presented in authors' edition. The authors are responsible for the content of their paper. The organizers, committees, reviewers, or editors reserve the right not to be responsible for the topicality, correctness, completeness or quality of the information provided.

2022 International Conference on Innovative Solutions in Software Engineering (ICISSE), Mykola Kuz and Mykola Kozlenko Eds., Ivano-Frankivsk, Ukraine: Vasyl Stefanyk Precarpathian National University, 2022, 317 p. [E-book]

ISBN 978-966-640-534-3

DOI 10.5281/zenodo.7502536

The high-quality refereed papers appearing in this book compose the proceedings of the 2022 International Conference on Innovative Solutions in Software Engineering (ICISSE). ICISSE is the forum organized by the Department of Information Technology of the Vasyl Stefanyk Precarpathian National University. ICISSE-2022 is held from November 29 to 30, 2022 in Ivano-Frankivsk, Ukraine. The event is intended to bring together researchers, scientists, and engineers to discuss experimental and theoretical results, mainly in the area of software engineering. The conference also covers topics related to computer science, computer engineering, systems analysis, cybersecurity, information systems and technology, industrial automation, electronics, metrology, micro and nanosystems, telecommunications, radio frequency engineering, IT entrepreneurship, and IT education. The current edition of the event is dedicated to the 50th birth anniversary of famous Ukrainian scientist working in information technology, Founder and first Head of the Department of Information Technology, Vice-Rector for Research of the university, Professor Dr. Pavlo Fedoruk, who passed away in 2013. The conference is included into the list of scientific conferences scheduled for 2022 by the Ministry of Education and Science of Ukraine.

Recommended paper citation template:

J. K. Author, "Decapitalized title of the paper," *2022 International Conference on Innovative Solutions in Software Engineering (ICISSE)*, Vasyl Stefanyk Precarpathian National University, Ivano-Frankivsk, Ukraine, Nov. 29-30, 2022, pp. xxx-xxx.

ISBN 978-966-640-534-3

© Copyright for the individual papers by the papers' authors, 2022
© Copyright for the book as a collection by its editors, 2022
© Vasyl Stefanyk Precarpathian National University, 2022

COMMITTEES

General Chair:

Prof. Dr. Mykola Kuz

Technical Program Committee Chair:

Assoc. Prof. PhD Mykola Kozlenko

Local Organizing Committee Chair:

Mariia Dutchak

Publication Chair:

Assoc. Prof. PhD Mykola Kozlenko

Preface

The 2022 International Conference on Innovative Solutions in Software Engineering dedicated to the 50th anniversary of an outstanding researcher and teacher, the founder of the Department of Information Technology at Vasyl Stefanyk Precarpathian University, Prof. Dr. Pavlo Fedoruk was held on November 29-30, 2022 in Ivano-Frankivsk, Ukraine.

Vasyl Stefanyk Precarpathian National University has more than 80 years of history. Its inspirational success story was written by people who devoted themselves to science, education, and culture. Among the prominent figures of the University, the name of the outstanding modern Ukrainian scientist Pavlo Fedoruk shone as a bright star. Now he would have celebrated his 50th birthday.

Pavlo Fedoruk graduated from Precarpathian University in 1994. Immediately after graduating, he was offered the position at the Information Department of the University. He proved himself to be a professional in the field of information technology, a talented person with progressive thinking. He was offered the position as a head of the Information Department in October 1994. Since 2002, Pavlo Fedoruk worked as a director of the Center for Information Technology. He also founded the Center of Distance Learning. While performing his official duties, he was simultaneously involved in research. He was a hard worker and independent thinking researcher who worked effectively on telecommunication technology research.

He studied as a PhD student in the period from 1999 to 2002. He received his PhD from the Institute of Mathematical Machines and Systems of the National Academy of Sciences of Ukraine. The topic of the PhD thesis was related to the systems of distance learning and knowledge assessment based on Internet technology. During his work at the University, Pavlo Fedoruk published more than 120 research papers and two monographs. He developed a course on Informatics for students of the Faculty of Economics and published the textbook "Computer for Economists". Also he developed courses on artificial intelligence and intelligent systems for students of the Faculty of Mathematics and Computer Science. In the period from 2006 to 2008, he was a scholarship holder of the Cabinet of Ministers of Ukraine for young scientists. In 2010, he was successfully awarded with the degree of Doctor of Technical Science. The topic of the thesis was related to adaptive systems of distance learning.

Since 2007, he was a member of the program committee and a session chairman at the international scientific and practical conferences "Education based on web technologies" and "Advanced computer technology in education". In 2007, he participated in the International Referee Committee as a representative from Ukraine. In 2008, he was an expert of the Scientific and Technical Council of the Ministry of Education and Science of Ukraine at the Informatics and Cybernetics section. He was a member of specialized academic councils at the Institute of Mathematical Machines and Systems and Ternopil National Economic University. He supervised four research projects in the field of distance learning. In 2009, he was awarded the Tersenov Regional Prize for Innovation from the Society of Inventors and Innovators of Ukraine. In 2010, he was recognized as a Famous Scientist of the Year for significant achievements in the field of research and education and for personal contribution. In 2011, he was awarded an honorary certificate from the Cabinet of Ministers of Ukraine for his significant personal contribution to science.

In 2012, he was promoted to the position of vice-rector for research of the University. His views were based on a broad knowledge and interests. He was a founder of many scientific, social, and entrepreneurial projects. He was one of the founders of the "Firtka" Local News Agency. He was also the President of the Ivano-Frankivsk Regional Martial Arts Federation.

The authors of this collection devoted their papers to the memory of Prof. Dr. Pavlo Fedoruk.

TABLE OF CONTENTS

Paper ID	Paper	Page
2	Moulfi Bouchra, Ferouani Souheyla, Ziani Kerarti Djalal, Marouf Fatima Zohra. New design of gold nano printed antenna with Gallium substrate for optical transmission	1
3	Zeba Shamsi, Subia Shamsi, Sekar K. A survey of the dark web	5
7	Наталія Полукетова, Анатолій Переверзев, Катерина Дереза. Моделювання ефективності дистанційної освіти із застосуванням агентного підходу	7
8	Oleh Zaritskyi. A practical approach to evaluating the ethics index of artificial intelligence systems	15
9	Богдан Савенко. Розподілені системи виявлення зловмисного програмного забезпечення	22
10	Олександр Мельниченко. Метод керування та самоорганізована система планування маршрутів БПЛА	26
11	Ольга Дмитрієва, Ілля Бутенко. Розробка веб-додатку для дистанційного навчання	31
12	Андрій Яворський, Максим Карпаш, Віталій Цих, Любомир Жовтуля. Пілотна система енергетичного моніторингу для навчальної будівлі університету	34
13	Iryna Tarasevich, Alena Tarasevich, Mykhailo Kuzmenko. Impact of Covid-19 on university education digitalization: a critical interpretive synthesis of tensions and barriers for automation field	43
14	Olga Solovei. Selection of data preprocessing technique for imbalanced dataset	48
15	Катерина Островська, Іван Стівченко, Сергій Черський. Дослідження можливості використання нейромережових технологій для вирішення задачі класифікації дефектів металопрокату.	51
16	Катерина Островська, Денис Печений. Дослідження методів на основі нейронних мереж для аналізу тональності корпусу текстів	55
17	Катерина Островська, Анастасія Петренко. Дослідження ефективності методів обробки зображень у системах розпізнання облич	59
18	Анастасія Нічепорук, Андрій Нічепорук, Андрій Казанцев, Олег Парніцький. MitM кібератаки на локально-керуючий рівень кіберфізичної системи розумних мереж енергозабезпечення: визначення доменної області та поведінка системи	63
19	Олександр Мельников, Вікторія Денисенко. Постановка задачі розробки програмного забезпечення для оцінювання лісистості та оброблення інформації про лісові насадження	67
20	Viktor Bilishchuk. Application of virtual simulation for device design	71
21	Iryna Vashchyshak, Serhii Vashchyshak. The web application of calculation of parameters for a small hydroelectric power plant	74
22	Olena Hladka, Ivan Karpovich, Daria Buryan. Modeling technologies for assessment of information security risks in e-commerce	78
24	Михайло Музикін, Галина Нестеренко, Олег Стрелко, Ніна Ключкова. Роль навігаційних систем в організації ефективної роботи автомобільного транспорту	82
25	Борис Незамай. Прогнозування якості програмного забезпечення	86
26	Oleg Zhulkovskiy, Inna Zhulkovska, Serhii Panteikov, Kyrylo Muzychka, Olena Panteikova, Yulia Pavlenko. Mathematical model for studying the thermal regime of the slag-splashing lance	90
27	Анастасія Чучвара, Ольга Чернуха, Юрій Білушак. Системний підхід до математичного опису дифузійних процесів у випадково неоднорідних двофазних структурах з кульовими включеннями	93
28	Oleh Pihnastyi, Maksym Sobol, Dmytro Yelchaninov. On the characteristics of the input material flow of the transport conveyor	102
29	Pavlo Katin, Oleksandr Pokhlyenko. Use of state pattern with the ability to dynamic add new states for software part of automatic cleaning-in-place stations for fermentation tanks	107
30	Petro Kostrobij, Vitalii Poloyi, Iryna Ryzha. Behaviour of SPP waves frequency spectrum for dielectric/metal/dielectric structures	115
32	Vadym Malinovskiy, Leonid Kupershtein, Vitaliy Lukichov. Risks assessment and approaches to creative of the reliable software modules for IoT devices	122
33	Михайло Крихівський, Любов Саманів. Математична модель інформаційного забезпечення екологічних служб	128
34	Roman Dorosh. Application of blockchain and the SimHash algorithm to detect Plagiarism	130

35	Василь Олійник, Андрій Подорожняк, Наталія Любченко. Аналіз та розпізнавання спаму в соціальних мережах та месенджерах	135
36	Олександр Жеребцов, Арсеній Рассохін. Проектування виконавчого пристрою системи контролю управління доступу на базі STM32F103	139
37	Юлія Казимиренко, Оксана Гайдаєнко, Владислав Принць, Анастасія Байталюк. Застосування нечітких алгоритмів в задачах розробки інформаційних систем для прогнозування результатів кіберспортивних змагань	147
38	Андрій Павлишин, Володимир Грига, Володимир Мандзюк. Реалізація алгоритму швидкого перетворення Фур'є на ПЛІС	149
39	Oleksandr Tsurkan, Yuliia Prokop. Automatic number plate recognition from videostream	157
40	Olena Oleinik, Ihor Michurin. Using Goroutine to accelerate SQL queries for PostgreSQL in AWS	160
41	Олександр Жеребцов, Антон Іжиков. Використання функціонального програмування у JavaScript та фреймворках	162
43	Денис Гобов. Контекст IT-проекту та проблеми бізнес-аналізу	169
45	Олександр Жеребцов, Марк Бороденко. Створення багатофункціонального локального веб-серверу на базі ESP8266	173
46	Roman Melnyk, Ruslan Tushnytskyu, Andriy Shpek. Selection of defective chains in a PCB image by thinning, flood-filling and overlaying	179
47	Віра Гарасимів. Автоматизоване тестування в Selenium: Page Object та Page Factory	184
48	Олександр Мельников, Дмитро Козуб. Постановка задачі розробки програмного забезпечення для оцінювання ефективності протиепідемічних заходів та прогнозування зміни відсотка інфікованих та перенесених хвороб у тяжкій формі	186
49	Юлія Сурганова, Юрій Міхлін. Нелінійні нормальні моди маятникової системи під дією магнітних збуджень	190
50	Ярослав Марчук. Дослідження використання паралельних технологій у високо-навантажувальних додатках	193
51	Ivan Rzhetskyi, Anastasiya Troynina. Garbage classification for mobile devices based on MobileNet	201
52	Наталія Іванічок, Олег Іванічок, Софія-Вікторія Склепова, Павло Колковський, Богдан Рачій, Роман Лісовський. Моделювання імпедансних спектрів, отриманих для системи пористий електрод/електроліт	205
53	Роман Базилевич, Андрій Франко. Класифікація структурних елементів програмного забезпечення у задачах генерації тестів	209
54	Роман Боднар. Пристрій для дослідження поверхневого натягу рідин методом газового струменя	212
56	Nataliia Liubchenko, Andrii Podorozhniak, Daniil Onishchenko. Development of forestry application using deep learning technology	216
57	Валентин Соболев, Андрій Подорожняк. Алгоритми пошуку шляху для безпілотних сільськогосподарських систем	218
58	Ольга Чорна, Данило Непорада, Олексій Чорний, Денис Калінін. Мобільний додаток «Система моніторингу асинхронних двигунів» на базі ОС Android	220
59	Тетяна Селівьорстова, Михайло Калюжний, Юрій Каліберда. Програмна реалізація сервісу агрегації даних для дослідження комерційних ринкових пропозицій автомобільного ринку України	224
60	Сергій Альошин, Олена Гайтан. Інтелектуальна технологія синтезу автоматизованої експертної системи керування підприємством	226
61	Mukhailo Kvasnii, Vladyslav Shevchuk, Halyna Klym. Intelligent system for collecting and analyzing of meteorological data based on robotic platform	232
63	Яна Лебеденко, Юрій Міхлін. Дослідження стаціонарного режиму та перехідного процесу в неідеальних системах	235
66	Вікторія Пахомова. Прогнозування часу доставки даних в ІТС залізничного транспорту з використанням нейронечіткої технології	238
68	Vasyl Franchuk, Nataliia Franchuk. Using Microsoft Family Safety for parents and children	243
69	Олена Смалько. Про сучасне різноманіття технологій та архітектурних рішень Інтернету речей	251
70	Vyacheslav Fomenko, Oksana Shovkopliias. Informational system of relief generation based on Perlín's noise mathematical algorithm	254

72	Ihor Polataiko. Software architecture design of software developed in student engineering teams	257
73	Валерій Ткачук, Олександр Кулик. Реалізація еволюційного алгоритму на реальному квантовому комп'ютері	262
75	Марія Дутчак, Микола Пікуляк. Оцінка якості методів та програмних засобів інтелектуальних адаптивних освітніх веб-систем самонавчання	266
76	Микола Пікуляк, Марія Дутчак. Програмні аспекти виявлення академічного плагіату	271
77	Mariia Slobodian, Mykola Kozlenko. Machine learning based animal emotion classification using audio signals	277
78	Yaroslav Drin, Iryna Drin, Roman Drin. The third initial-boundary value problem for the nonlinear equation of heat conduction with deviation of the argument	282
79	Владислав Гой. Дослідження масштабування та розробка рекомендаційної системи методами глибокого навчання та технологій великих даних	286
80	Дмитро Юрчук, Іван Савка. Дослідження шкідливого програмного забезпечення	291
81	Микола Кузь, Ігор Лазарович, Руслан Мішагін. Напрямки удосконалення міжнародних стандартів у сфері оцінки якості програмних продуктів ISO/IEC 9126 та ISO/IEC 14598	294
82	Vasyl Ivanochko. Analysis of methods and existing software implementations for automatic resume generation	298
83	Yuliia Dudii. Overview of methods and existing solutions for task order generation	306
84	Valentyna Yakovyn, Taras Kriukov. Object detection in mobile robotics	310
85	Mykola Kozlenko. Supervised machine learning based signal demodulation in chaotic communications	313

New design of gold Nano printed antenna with Gallium substrate for optical transmission

Moulfi Bouchra, Ferouani Souheyla, Ziani Kerarti Djalal, Marouf Fatima zohra

*Department of Electronic and Telecommunications SSL Laboratory
Univ Belhadj Bouchaib (of Ain Temouchent)
Temouchent, Algeria*

Abstract—IN this paper we will simulate a circular patch antenna in the terahertz band for WBAN (wireless body area network) . the material chosen for the patch and the ground plane is gold and for the substrate is Gallium ($\epsilon_r=12.94$). the result of simulation is very satisfactory, we have a return lost -44 dB in frequency 3.85 THZ and Radiation pattern of 3.78 dB. The bandwidth obtained is more than 650 Ghz .

Keywords— *patch antenna, WBAN, return lost, Radiation pattern.*

I. INTRODUCTION

THZ antennas have a great interest by researchers due to their amazing characteristics, several works have been done in this frequency spectrum, antennas based on new materials [1][2], lens based antennas [3], gold based antennas [4]. Terahertz optical antennas can be used in sensitive areas such as medicine which requires very small antennas for easy integration into the human body, omnidirectional radiation which does not hurt the human body [5][6].

The WBAN (wireless body area network) uses the optical spectra pocket to the microwave [0.5-4] THZ[7][5] to measure temperature, oxygen, ECG, EEG etc[8]. of patients who must remain under medical supervision or in certain sports[9]. There are three types of WBN: WBN IN-Body, ON-Body, OFF-Body[10]. This type of antenna has been used in the last years after the apparition of contagious diseases such as COVID. This led to the risk of staying in the hospital especially for people with conical diseases.

In this work we will design with software CST Microwave a circular WBAN patch antenna abraded with gold at the frequency 3.85 THZ, using Gallium($\epsilon_r=12.94$) as a substrate material between the ground plane and the patch with a thickness of 5.4 μ m. The technique of intergrations of slot on the ground plane and interred for the adaptation of the antenna.

II. PROPOSED ANTENNA

In this part we will design a circular patch antenna at the frequency 3.85 THZ, fed by a micro rubun line with an impedance of about 50 Ω , the material used for the patch and the ground plane is gold and the substrate used is Gallime ($\epsilon_r=12.94$). The dimensions of the antenna and generally depend on the operating frequency they have calculated with the following formula:

$$F = \frac{8.791 \cdot 10^9}{f r \cdot 10^{12} \cdot (\epsilon_{ps})^{0.5}} * 10^4 \quad \text{Eq. 1}$$

$$a = \left(\frac{F}{\left(1 + \left(\frac{2 \cdot h \cdot 10^{-4}}{\pi \cdot \epsilon_{ps} \cdot F}\right) \cdot \ln\left(\left(\frac{\pi \cdot F}{2 \cdot h \cdot 10^{-4}}\right) + 1.7726\right)\right)^{0.5}} \right) \quad \text{Eq. 2}$$

$$ae = a * \left(1 + \left(2 * \frac{h}{3.14 * a * \epsilon_{ps}}\right) * \ln\left(\frac{3.14 * a}{2 * h}\right) + 1.7726\right)^{0.5} \quad \text{Eq. 3}$$

$$Lg = ll + 2 * ae + 6 * h \quad \text{Eq. 4}$$

$$Wg = 2 * ae + 6 * h \quad \text{Eq. 5}$$

Table 1 shows the dimensions of the proposed antenna and figure 02 shows the proposed circular antenna.

Antenna parameters	Dimension in (μm)
a	6.1
h	5.4
wl	3.93
Ll	19.26
wg	63
Lg	22

TABLE 1. size of the proposed antenna

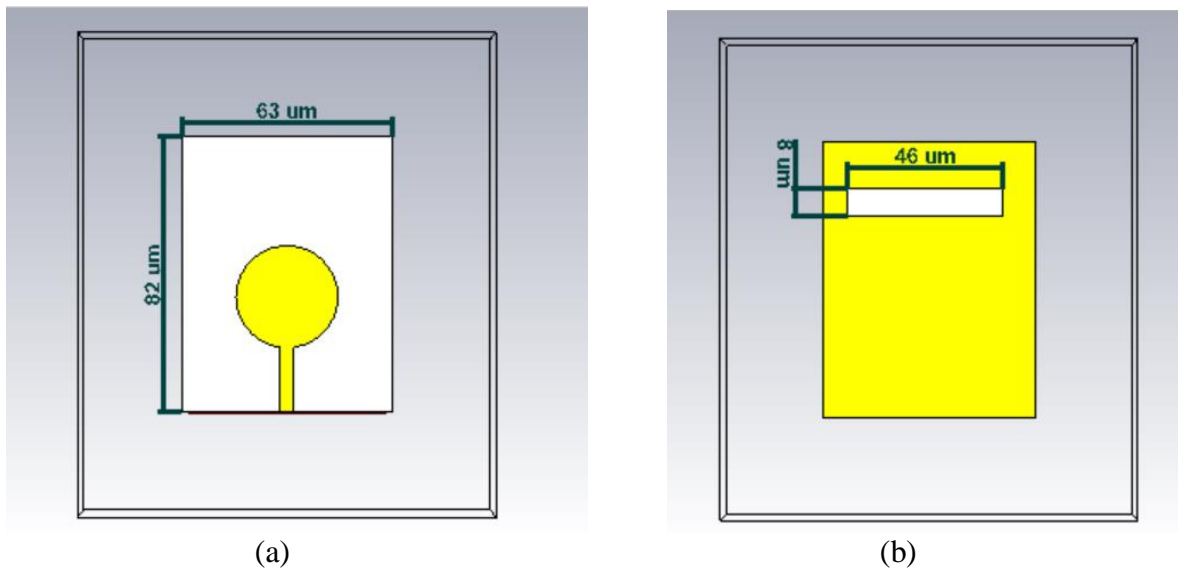


Fig. 1 .circular proposed graphene patch antenna

(a)front view, (b) reare view.

III. RESULTS

A. Return loss S11

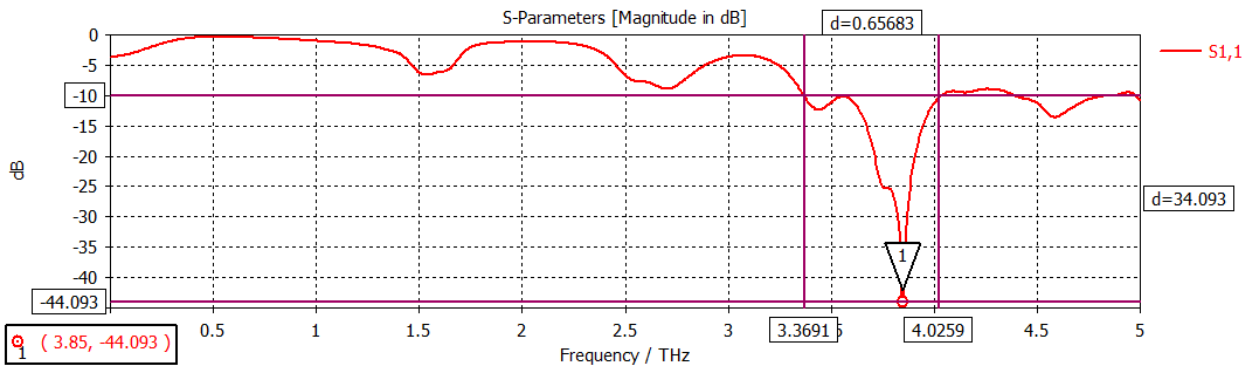


Fig. 2. reflection coefficient S11

Figure 2 shows the simulation result of the reflection coefficient of the proposed antenna. The Gallium ($r=12.94$) gives a very good result at the frequency 3.85 THz, we have an S11 of -44.093 dB and a bandwidth of 656.83 GHz.

B. Radiation pattern (Gain)

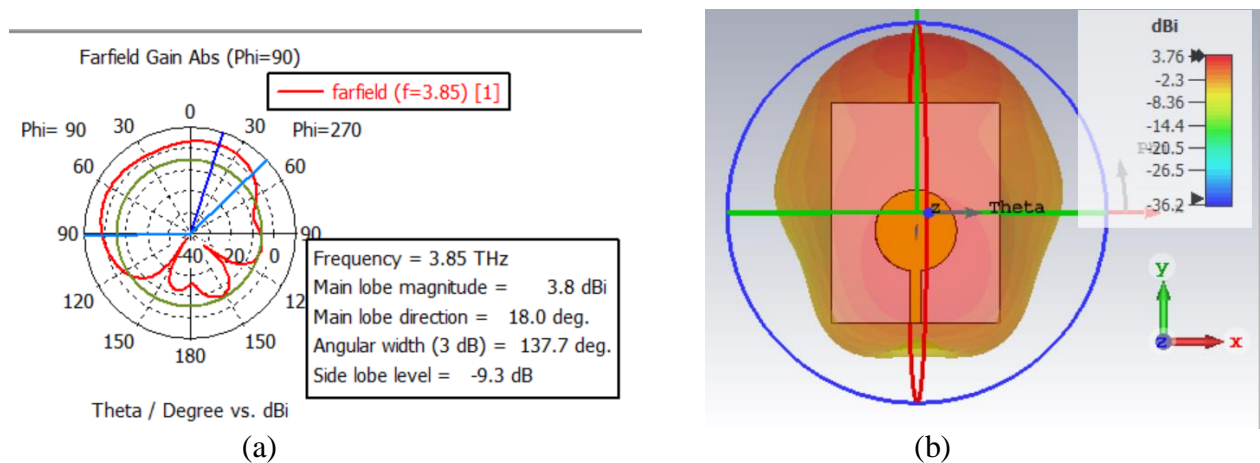


Fig. 3.(a) polar Radiation pattern (b) 3D Radiation pattern of proposed antenna

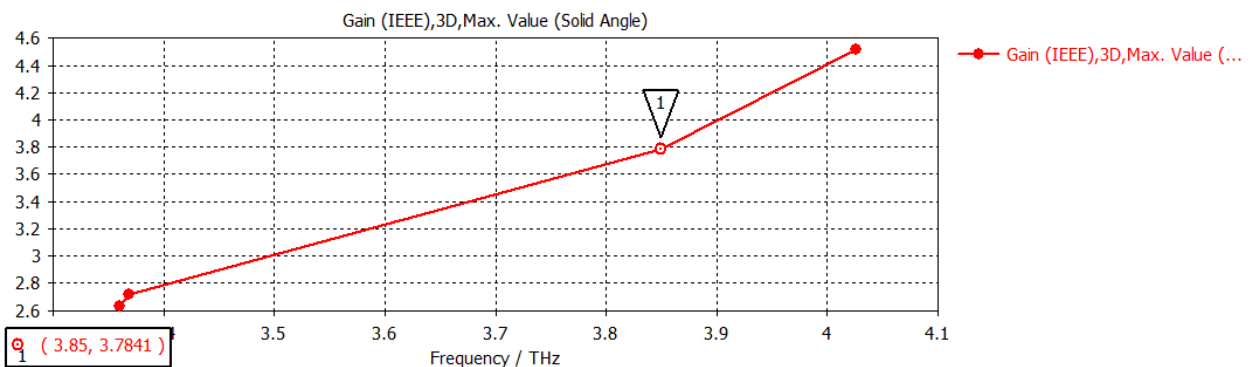


Fig. 4 Gain of proposed antenna

Figure 3 shows the gain of the proposed antenna in polar and 3D and figure 4 shows the gain in the whole bandwidth. These figures show that the proposed antenna has an omnidirectional gain exceeding 2.6 dB in the whole frequency band, so this antenna can be used in the WBAN which requires an omnidirectional gain exceeding 2 dB.

C. VSWR

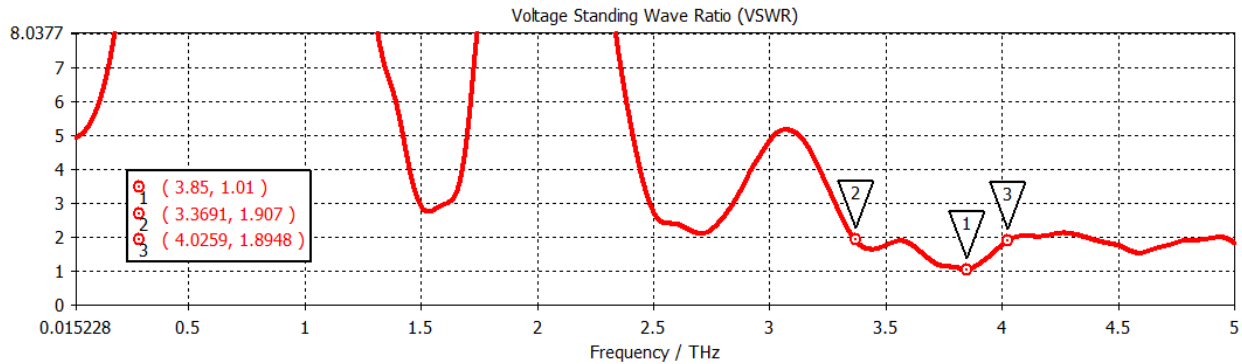


Fig. 5. VSWR

The VSWR obtained is ≤ 2 as shown in figure 5

IV. CONCLUSION

In this work we will design a circular antenna for WBAN applications using CST software, the result of simulation and be satisfactory for wireless body area network. The proposed antenna has a gain of 3.78 dB and a reflection coefficient S_{11} -44 dB in a frequency of 3.85 THz.

V. ACKNOWLEDGMENT

The authors thank laboratory of the intelligent structure of the university of Ain Temouchent (SSL) for their technical assistance in the actual research.

VI. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] M. Bouchra, F. Souheyla, K. DjalalZiani, and W. Moulessehoul, "Design of a Novel Nanometric Graphene Pentagone Patch Antenna Array for Terahertz Transmission," *IEEE*, pp. 1–5, doi: 10.1109/ICISAT54145.2021.9678495.
- [2] P. Ram and C. Singh, "A Novel Graphene Conductive Ink Based Circular Patch Antenna for 2.4 GHz Application," *Wirel. Pers. Commun.*, vol. 116, no. 4, pp. 3101–3108, 2021, doi: 10.1007/s11277-020-07837-7.
- [3] S. Ebrahimi, "Microfluidic as liquid lens for THz reconfigurable antenna and gain enhancement with sensing application," *Optik (Stuttg.)*, p. 167705, 2021, doi: 10.1016/j.ijleo.2021.167705.
- [4] V. Das and S. Rawat, "Modified rectangular planar antenna with stubs and defected ground structure for THz applications," *Optik (Stuttg.)*, vol. 242, no. April, p. 167292, 2021, doi: 10.1016/j.ijleo.2021.167292.
- [5] E. P. J. Parrott, Y. Sun, and E. Pickwell-Macpherson, "Terahertz spectroscopy: Its future role in medical diagnoses," *J. Mol. Struct.*, vol. 1006, no. 1–3, pp. 66–76, 2011, doi: 10.1016/j.molstruc.2011.05.048.
- [6] Q. Rubani, S. H. Gupta, and A. Kumar, "Design and analysis of circular patch antenna for WBAN at terahertz frequency," *Optik (Stuttg.)*, vol. 185, pp. 529–536, 2019, doi: 10.1016/j.ijleo.2019.03.142.
- [7] D. Arnone, *Applications of Terahertz (THz) Technology to Medical Imaging*, vol. 3828, no. 1, 1999.
- [8] S. N. Mahmood *et al.*, "Full ground ultra-wideband wearable textile antenna for breast cancer and wireless body area network applications," *Micromachines*, vol. 12, no. 3, pp. 1–16, 2021, doi: 10.3390/mi12030322.
- [9] M. O. Santos, S. M. M. Faria, and T. R. Fernandes, "Real time underwater radio communications in swimming training using antenna diversity," *2021 Telecoms Conf. ConfTELE 2021*, 2021, doi: 10.1109/ConfTELE50222.2021.9435592.
- [10] M. Kumar, S. Goel, A. Rajawat, and S. H. Gupta, "Design of optical antenna operating at Terahertz frequency for In-Vivo cancer detection," *Optik (Stuttg.)*, vol. 216, p. 164910, 2020, doi: 10.1016/j.ijleo.2020.164910.

A Survey of the Dark Web

Zeba Shamsi¹, Subia Shamsi², Sekar K³

¹Department of Computer Science and Engineering, Al-Falah University, Faridabad- Haryana, India

²Department of Computer Science and Engineering, National Institute of Technology Silchar, Assam, India

Abstract— The dark web is a hidden network of websites that can only be viewed by using a particular web browser. It protects users and anonymity through technologies such as traffic encryption, obfuscation, and forwarding. This study looks at how the dark web is being used, focusing on criminality and how law enforcement acts as an opponent. The review explains these secret areas, including their history, the activities they host such as cybercrime the nature of the attention they attract, and the tactics used by law enforcement to try to defeat their goal.

Keywords—Dark Web, Cybercrime, Law enforcement.

I. INTRODUCTION

The Internet provides a platform for an information system, known as the World Wide Web, where a global community can exchange information. Within this community exists the Dark Web or Darknet, an environment that affords its users anonymity, making attribution for activities challenging by encrypting and routing users' traffic via multiple nodes. This paper structured as first introduction, related work and search methodology of the area of Dark Web. Then discuss the literature of the phenomenon of dark web, the significant roles the dark web plays, and discuss legal and societal concerns. Lastly, we present a future research agenda and conclusions of the paper.

II. RELATED WORK

As the internet has grown in popularity, so has the demand for users to be able to function anonymously. Onion Router (Tor) systems, crypto currency, and other dark web applications have grown in popularity, turning the internet's ubiquitous promise into a shell game between the good and the bad [1]. The author proposed the Tor use motive model, a two-dimensional categorization system built primarily for use in law enforcement [2]. When it comes to law enforcement, the majority of vendors on dark web platforms have a record in the drugs trade. People who sell illegal goods online do so because they anticipate more profits, lower risks of being caught by law enforcement agencies, and less chances of physical harm or financial loss [3].

III. RESEARCH AGENDA

This section discusses the research gaps identified in the area of current Dark Web literature. We concluded each area of the set of one or more future research questions to set up the research agenda. These research questions aim to fill in the gaps in current dark web research. Cybercriminals have benefited from hidden services by purchasing vulnerabilities on black markets; interacting with other criminals on forums; running command-and-control (C2) servers for botnets and exfiltration malware, and employing professional criminals or their talents 'as-a-service.' The dark web's anonymity feature allows thieves to carry out their operations [4]. There are many different forms of cybercrime and actors, ranging from inexperienced script-kiddies who utilize widely accessible tools and vulnerabilities to highly organized Advanced Persistent Threat actors who have been linked to substantial incidents of espionage against huge organizations and nation-states. These assaults are aimed at the digital assets of businesses. There was no extensive literature specifically addressing a relationship between the dark web and the APT phenomenon. In future more related researches would be done to reduce or better protect against cyber-attacks.

As technology advances, tracking, identifying, and shutting down sites that allow illegal activity becomes more difficult. The dark web allows communication between people to become more

decentralized and thus harder to regulate. Law enforcement concentrates its efforts on deanonymizing hidden services, marketplace operators, site administrators, and anybody else engaged in facilitating unlawful activity.

IV. RESULTS

The Dark Web is increasingly essential in cybercrime, and law enforcement's role in preventing unlawful conduct on the dark web was the subject of three study topics. While online marketplaces for illicit narcotics remain popular, they are only one of several options accessible, which range from software vulnerabilities to extremist recruiting.

V. DISCUSSION

The dark web is a characteristic of a unique distributed system that exists above the global internet and claims to be a sanctuary from prying eyes, mostly law enforcement. Technological advancements have enabled the dark web to live and thrive for longer than ever before.

VI. FUTURE RESEARCH

Future this debate will be over as technology advances to the point where privacy is not at the mercy of governments but in the hands of users and private corporations. The discussion on this topic then no longer remains in the domain of technology but is one that needs interdisciplinary contemplation by experts in areas of psychology, sociology, law, and others.

VII. CONCLUSION

This paper conducts a systematic survey into the dark web's roles in modern digital society, its enablement of cybercrime, and its relationship with law enforcement and society. Surveying this topic was challenging as the information available and research conducted on the topic is relatively limited due to its secretive and non-identifying nature. Law enforcement continues in its attempts to close markets responsible for the sale of illicit items, especially recreational drugs. On the other hand, the infrastructure responds to these incidents and attempts to mitigate the vulnerabilities exploited by the authorities. As a result, the Tor protocol is constantly updated to address protocol vulnerabilities; in the debate over privacy versus security, the technology factor is weighing in more heavily than ever before, as it becomes a matter of not only legality but also technical capability to monitor and conduct surveillance on people.

VIII. ACKNOWLEDGMENT

The authors gratefully acknowledge the faculty members from the computer science and engineering department, National Institute of Technology Silchar, Assam India.

IX. DISCLOSURES

The authors have no conflicts of interests to disclose.

REFERENCES

- [1] Nina Sunde, and Graeme Horsman, "Part 2: The Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations", *Forensic Science International: Digital Investigation*, Volume 36, 301074, ISSN 2666-2817, (2021). <https://doi.org/10.1016/j.fsidi.2020.301074>.
- [2] Janis Dalins, Campbell Wilson, Mark Carman, "Criminal motivation on the dark web: A categorisation model for law enforcement", *Digital Investigation*, Volume 24, Pages 62-71, ISSN 1742-2876, (2018). <https://doi.org/10.1016/j.diin.2017.12.003>.
- [3] James Martin, Rasmus Munksgaard, Ross Coomber, Jakob Demant, Monica J Barratt, Selling Drugs on Darkweb Cryptomarkets: Differentiated Pathways, Risks and Rewards, *The British Journal of Criminology*, Volume 60, Issue 3, Pages 559-578, (2020). <https://doi.org/10.1093/bjc/azz075>
- [4] Ransomware and dark web operations disrupted, and arrests made following major law-enforcement actions around the world, *Computer Fraud & Security*, Volume 2021, Issue 11, ages 1-3, ISSN 1361-3723, (2021). [https://doi.org/10.1016/S1361-3723\(21\)00111-1](https://doi.org/10.1016/S1361-3723(21)00111-1).

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

Моделювання ефективності дистанційної освіти із застосуванням агентного підходу

Наталія Полуектова, Анатолій Переверзєв, Катерина Дереза

*Кафедра інформаційних технологій
Запорізький інститут економіки та інформаційних технологій
Запоріжжя, Україна*

Анотація— У роботі представлено аналіз сучасних підходів до оцінки ефективності сучасних освітніх систем та технологій та запропоновано агентну модель для вивчення взаємодії та оцінки ефективності акторів у таких системах. Аналіз показав наявність проблем, пов'язаних із складністю освітніх систем, появою нових типів мережевих структур та нових вимог до ефективності освіти. Розроблена агентна модель враховує нові типи взаємодії між учасниками системи вищої освіти – не лише між викладачем та учнями, а й між елементами освітнього контенту та між студентами. Агентний підхід використовує внутрішні поведінкові фактори дійових осіб та зміни їхнього стану при зміні параметрів середовища. Параметри моделі також враховують вплив нових інформаційних та комунікаційних технологій на швидкість та якість навчання. Ряд стохастичних експериментів із моделлю проведено з використанням системи моделювання AnyLogic. Результати експериментів дозволяють оцінити вплив складності змісту, здібностей учнів, ступеня їх взаємодопомоги та інших факторів на показники ефективності в класичному і дистанційному навчанні. Модель може стати основою вивчення проблем ефективності сучасних освітніх систем і технологій.

Ключові слова— *Агентне моделювання, Дистанційне навчання, Ефективність ІТ освіти*

I. ВСТУП

Бурхливий розвиток інформаційно-комунікаційних технологій у світі за останні десятиліття спричинив важливі зміни в усіх сферах життя та діяльності людини. На авансцену цих змін вийшли процеси, пов'язані з наукою та освітою. Традиційні навчальні заклади мають дуже швидко реагувати на нові виклики. Зростання глобалізації, швидке старіння знань і досвіду, поява нових міждисциплінарних напрямів досліджень, потреба в більш тісній взаємодії між науковим співтовариством і виробничими системами є одними з цих викликів. Таким чином, навчальні заклади стають одним із гравців у значно розширеному процесі виробництва знань. Наслідком цих процесів є розширення широкого розмаїття мереж співпраці для генерації знань, зростаючої конкуренції за ресурси, контракти та студентів. Процес виробництва нових знань потребує вдосконалення.

Ефективність будь-якої системи професійної освіти оцінюється на основі двох основних підходів: цільового або технологічного. Цільова ефективність - це ступінь відповідності функціональних параметрів досліджуваної системи цільовим значенням цих параметрів. У нашому випадку розглядається ступінь досягнення освітніх цілей. З іншого боку, система освіти виступає як якась виробнича система, яка споживає ресурси і дає результати. Їх оцінка базується на фінансових доходах освітньої організації. Це дозволяє говорити про технологічну (ресурсну) ефективність.

Автори огляду про існуючі підходи до оцінки ефективності вищої освіти підтвердили цю тезу [1]. Вони розглядають підхід на основі ресурсів і підхід на основі цінностей.

Ресурсний підхід оцінює продуктивність навчального закладу як ступінь досягнення ним своїх цілей з мінімальною кількістю ресурсів. У вищій освіті продуктивність часто є оцінкою кількості студентів на одного викладача або кількості опублікованих журнальних статей на одного дослідника. Використання такого підходу викликає проблему визначення цілей, з 2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

одного боку, та оцінки вартості ресурсів, з іншого.

Підхід, заснований на цінностях, фокусується на результатах, досягнутих для зацікавлених сторін, включаючи студентів, роботодавців, місцеву громаду та суспільство в цілому. У цьому випадку потрібно враховувати як кількісні, так і якісні показники. Усі три елементи – зменшення вартості ресурсів, ефективність і дієвість – однаково важливі в контексті сучасної освіти.

Зауважимо, що проблема оцінки ефективності систем освіти в останні роки набула ще більшої актуальності, що пов'язано з тенденцією, яка спостерігається в більшості країн Східної Європи, а також у Великобританії, Австрії та деяких інших країнах. У більшості країн попит на спеціальну освіту зростає. Однак уряди намагаються підтримувати той самий рівень фінансування. Це призводить до необхідності економії грошей і пошуку нових шляхів підвищення ефективності [2].

Впровадження різноманітних форм дистанційного навчання стає одним із найважливіших напрямків у розвитку освітніх технологій. Вони дозволяють організаціям, що надають послуги освіти, вирішувати нові виклики. Актуальність дослідження проблеми якості та ефективності дистанційного навчання неочікувано зросла у зв'язку з епідемією коронавірусу, коли більшість вищих навчальних закладів світу потребували швидкого впровадження нових форм і технологій навчання, і, ці процеси ще поглиблюються за умов війни.

В роботах [3,4] представлені офіційні визначення дистанційної освіти. Ці дослідження підкреслюють наступні характеристики, які відрізняють системи дистанційного навчання від традиційних освітніх систем:

1. фізичне розділення вчителя та учня в процесі навчання;
2. індивідуальна двостороння взаємодія вчителя і учня;
3. вибір викладачем різноманітних методів і технологій, що полегшують стосунки зі студентом під час навчання;
4. наявність формалізованої навчально-методичної процедури, яка визначається навчальним закладом, а не викладачем.

Деякі переваги використання таких технологій для студентів очевидні: можливість вибору та дистанційного навчання за програмами кращих навчальних закладів; можливість продовжувати вдосконалювати свої знання та вміння протягом життя; повністю змінити професію, не змінюючи звичного способу життя [5]. Важливим фактором також є, як правило, менша вартість таких форм навчання для студентів.

Навчальні заклади також можуть отримати певні переваги від впровадження систем дистанційного навчання. Це пов'язано зі здешевленням денної форми навчання, залученням студентів і викладачів з інших регіонів і країн.

Важливо визначити суттєві відмінності, які можуть вплинути на ефективність навчання при впровадженні методів і технологій дистанційного навчання.

Деякі дослідження останніх років показують, що ефективність систем дистанційного навчання залежить не стільки від технологій, скільки від стратегій і стандартів забезпечення якості цифрових перетворень, прийнятих організацією [6].

Робота [7] довела, що в системах дистанційної освіти необхідно вивчати взаємодію кількох груп акторів.

1. Студент – контент. Студент самостійно вивчає матеріал курсів. Важливою особливістю та перевагою систем дистанційного навчання є можливість використання різноманітних джерел та форм подання такого контенту.

2. Студент - викладач. Індивідуальна взаємодія викладача з кожним студентом є найважливішою умовою успішності дистанційного навчання. Перевагою дистанційної системи є можливість використання різноманітних засобів зв'язку, іноді підвищується рівень доступності викладача.

3. Студент - студент. Групове спілкування підвищує мотивацію та є важливим чинником підвищення ефективності навчального процесу, а сучасні засоби спілкування, такі як соціальні мережі, наприклад, дозволяють досить ефективно організувати такі процеси.

4. Контент - контент. Сучасні інтелектуальні технології дозволяють автоматично оновлювати та вдосконалювати зміст електронних курсів, що, безумовно, сприяє підвищенню ефективності процесу, але потребує досить значних витрат на організацію та впровадження інтелектуальних технологій.

5. Керівник курсу - контент. Процес планування навчального курсу відіграє важливу роль у забезпеченні його ефективності. Керівник повинен стежити за своєчасним оновленням матеріалу, враховувати нові звернення студентів та інших зацікавлених осіб.

6. Викладач - викладач. Творче та професійне співтовариство викладачів дозволяє більш ефективно розробляти та впроваджувати нові методи та технології навчання, що є дуже важливим в умовах стрімкого розвитку ІКТ.

7. Студент - технічний інструктор, викладачі - технічний інструктор. У складних і добре організованих системах дистанційного навчання доступність технічної допомоги студентам і викладачам стає одним із найважливіших факторів підвищення ефективності.

У цій статті ми розглядаємо деякі аспекти поведінки цих акторів у системах освіти на основі теорії та практики агентного моделювання, яке є частиною сучасних методів моделювання. Крім того, ми спробуємо зрозуміти, як у такому підході можна врахувати зміни, пов'язані з використанням дистанційного навчання.

II. МЕТОДИ

Методи імітаційного моделювання активно використовують у задачах дослідження систем зі складною структурою, і забезпечують дослідження будь-яких складних і динамічних процесів з великою кількістю елементів. Такою складною системою є система сучасної освіти.

Усі методи імітаційного моделювання в даний час поділені на методи системної динаміки, дискретно-подійні методи та агентні методи [8]. У роботі [9] наводиться обґрунтування того, що саме агентний підхід найкраще підходить для вивчення процесів у системі професійної освіти, оскільки дозволяє дослідити причинно-наслідкові зв'язки, що виникають у такому багатоагентному середовищі.

У найбільш загальному вигляді агентний підхід може бути описаний марківським процесом прийняття рішень (MDP), де агент приймає стан навколишнього середовища в якості вхідних даних і генерує дію як вихідні дані, чим надалі впливає на навколишнє середовище. [10].

В основі представленої роботи лежить підхід, запропонований у [11], який описує систему взаємодії двох агентів – вчителя та учня в межах електронної класної кімнати та демонструє як різні набори правил взаємодії дозволяють змінити ефективність навчання. Цей підхід використовує теоретичну модель гри, що ґрунтується на дилемі Ітеративного ув'язненого [12] і моделі MetaGame of Learning (MGL).

Цей підхід у контексті вищої освіти може бути описаний наступним чином: викладач пропонує просте чи складне завдання, з яким студент може впоратися швидше чи повільніше. В результаті такої взаємодії викладач та студент прагнуть досягти основної мети освіти – отримання нових знань, якщо завдання досить складне, а студент досить старанно займається. Інакше ефективність системи знижується. При цьому в моделі може бути врахований і різний рівень здібностей студентів.

В оригінальній моделі описується гра двох агентів: викладача та студента. Однак, для вивчення впливу дистанційних технологій на систему освіти має бути побудована система взаємодії всіх вищеперелічених пар агентів та враховані нові параметри, що виникають у системі навчання (середовищі) при впровадженні нових технологій дистанційної взаємодії.

Успіх сучасних агентних моделей багато в чому зумовлений застосуванням адекватних програмних інструментів. В даному випадку була використана система AnyLogic, яка, на наш погляд, має найбільш розвинені на сьогодні засоби візуалізації імітаційних моделей та організації експериментів різних типів.

Запропонована агентна модель розвивалася в такий спосіб.

На першому етапі було побудовано модель, що включає такі типи агентів:

- один викладач;
- популяція з 10 студентів;
- порожня популяція агентів, що включає два типи інформації (дочірні агенти) – нова інформація, яка описує матеріали одного з тематичних модулів деякого курсу, що вивчаються, і, уточнююча інформація, яка дозволяє описати взаємодію викладача та студентів у процесі вивчення тематичного модуля у формі питань і відповідей;
- популяція запитів, необхідна для того, щоб студент міг генерувати два типи запитів – запит на отримання нового блоку інформації та запит на отримання відповіді викладача в процесі консультування.

Поведінка кожного типу агентів визначається набором параметрів і діаграмами станів, в яких може перебувати кожен з агентів.

У цьому варіанті моделі викладач не є активним агентом і служить лише для генерації та розподілу інформації.

На рис. 1 представлена діаграма станів агента Студент. Кожен із агентів типу Студент може перебувати в одному з трьох станів: самостійне вивчення матеріалів, консультування та очікування нового матеріалу. Перехід від стану самостійного вивчення до стану консультування виконується з певною частотою, яка задається параметром, що відображає час очікування відповіді викладача. Модель містить також імовірнісні характеристики переходу на основі зміни параметра часу очікування за експоненційним законом розподілу. А перехід у стан очікування нового матеріалу відбувається із частотою, визначеною параметром середнього часу вивчення окремого тематичного модуля. Повернення до стану самостійного вивчення відбувається при прибутті агента Інформація, що символізує відповідь на запитання або нову інформацію відповідно. В даному випадку студент перебуває у стані консультування в очікуванні відповіді на своє запитання.

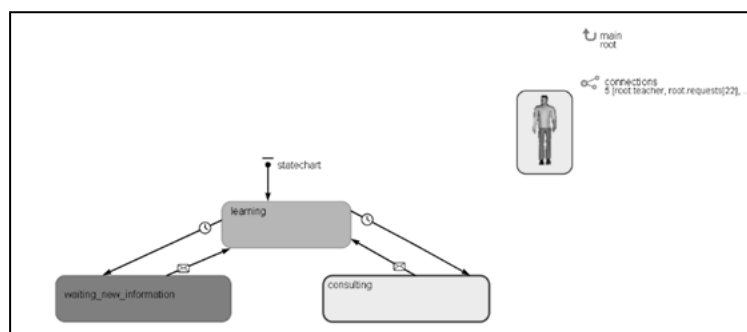


Рис. 1. Діаграма станів агента Студент

На рис. 2 представлена діаграма станів агента Нова Інформація.

З малюнка 2 зрозуміло, що нова інформація може перебувати у чотирьох станах – готуватися викладачем, передаватися студенту, вивчатися студентом, передаватися викладачеві як повідомлення про закінчення вивчення даного тематичного модуля. В даному випадку – інформація вивчається студентом.

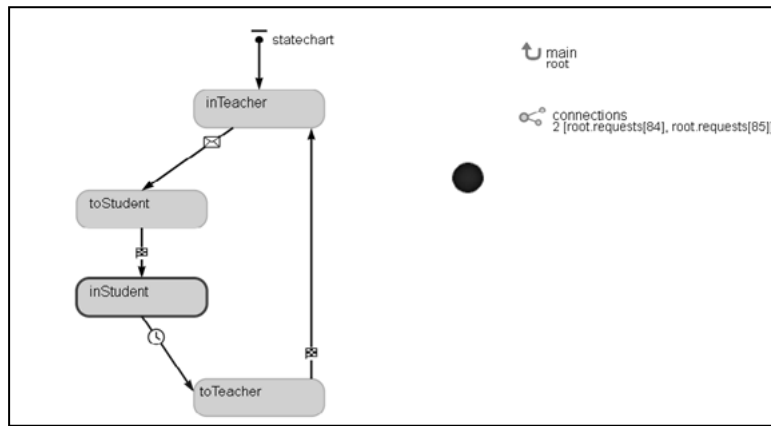


Рис. 2. Діаграма станів для агента Нова Інформація.

Результатом побудови першого варіанта моделі стала можливість оцінки ефективності вивчення матеріалу окремим студентом i , розрахована як:

$$E_i = 1 - (\sum_{j=1}^m (e^{\frac{1}{t_j^w + t_j^c}})) / t_{mod} \quad (1)$$

де t_j^w модельний час очікування відповіді викладача на j -е питання студента, t_j^c – модельний час безпосереднього консультування, t_{mod} – середній час вивчення модуля, m – кількість питань, які задає один студент.

Ще одним параметром, що оцінюється, стала ефективність системи в цілому, яка розраховується як середнє значення показника індивідуальної ефективності за всіма студентами:

$$E = \frac{\sum_{i=1}^n E_i}{n} \quad (2)$$

На наступному етапі до моделі було додано кілька параметрів, що дозволяють оцінювати вплив складності інформації та здібностей студентів на час, що витрачається на вивчення тематичного модуля. Для цього використовувався параметр c (complexity), який характеризував складність кожного екземпляра популяції Нової інформації та параметр a (ability) який дозволяв моделювати здібності кожного екземпляра із популяції Student. Таким чином з'являється можливість оцінювати зміни показників ефективності за рахунок коригування параметра час консультування з рівняння 1 наступним чином:

$$(t_j^c)' = (t_j^c) * c/a \quad (3)$$

Пізніше модель була модифікована, щоб дозволити врахувати можливість взаємодії між студентами в процесі навчання, їхньої допомоги один одному і, таким чином, прискорення переходу від стану самостійного вивчення матеріалу до стану очікування нового матеріалу. Ці зміни відбулися шляхом позначення двох додаткових переходів на діаграмі стану студента (рис. 3).

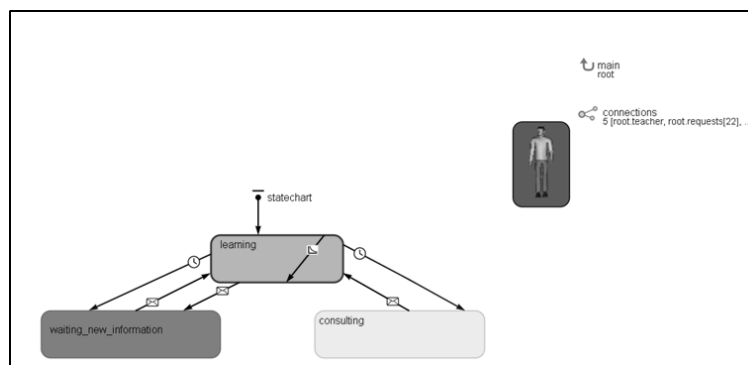


Рис. 3. Зміни на діаграмі стану студента, пов'язані з моделюванням взаємодопомоги.

З отриманою моделлю ми провели кілька серій експериментів. Це виявило деякі закономірності поведінки агентів у класичній системі навчання. Деякі параметри моделі явно змінюються при використанні методів дистанційного навчання. Це також дало змогу вивчити зміни показників успішності під час переходу на дистанційне навчання.

III. РЕЗУЛЬТАТИ

Основні результати проведених експериментів наведено в таблицях 1 і 2. Як базу для порівняння було прийнято показники, які отримані під час навчання студентів спеціальності «Програмна інженерія» щодо дисципліни «Web-програмування». В групі 3 курсу навчається 10 студентів, дисципліна поділена на 5 тематичних модулів.

З використанням вбудованих можливостей системи Anylogic було проведено вісім експериментів варіювання параметрів методом Монте-Карло. В результаті на підставі значень вихідних параметрів, що стохастично змінюються, були отримані значення обчислених показників. В таблиці 1 за розрахунками, проведеними на підставі отриманих даних у цій групі, були отримані результати, що імітують класичне (не дистанційне) навчання.

Таблиця 1. Результати експериментів з моделлю для класичної форми навчання

	Базовий варіант	1 експери мент	2 експери мент	3 експери мент	4 експер имент
Вхідні параметри					
Кількість контактів одного студента	0	0	0	0	5
Середній рівень складності матеріалу	0,5	0,5	0,8	0,5	0,5
Середній рівень здібностей студента	0,5	0,5	0,5	0,8	0,5
Вихідні параметри					
Загальний час консультування (год)	36	38	43	40	32
Індивідуальна ефективність	0,85	0,88	0,76	0,84	0,83
Загальна ефективність системи	0,8	0,76	0,68	0,73	0,85

В таблиці 2 наведені результати імітаційних експериментів, щодо дистанційного навчання.

Таблиця 2. Результати експериментів з моделлю дистанційної форми навчання

	Базовий варіант	5 експери мент	6 експери мент	7 експери мент	8 експер имент
Вхідні параметри					
Кількість студентів	10	10	10	20	20
Кількість модулів	5	5	5	5	10
Середня кількість питань при вивченні модуля для кожного студента	10	10	10	7	7
Середній час очікування відповіді на запитання (год)	0,2	0,2	0,3	0,3	0,3
Середній час консультації з одного питання (год)	0,5	0,5	0,5	0,5	0,5

Кількість контактів одного студента серед студентів	0	0	0	0	5
Середній коефіцієнт складності матеріалу	0,5	0,5	0,8	0,5	0,5
Середній коефіцієнт здібностей студента	0,5	0,5	0,5	0,8	0,8
Вихідні параметри					
Загальний час консультування	46	43	48	35	31
Індивідуальна ефективність	0,61	0,58	0,55	0,82	0,85
Загальна ефективність системи	0,59	0,57	0,56	0,8	0,89

IV. ОБГОВОРЕННЯ

Перший експеримент повторює дані з базового прикладу. Отримані результати свідчать про придатність даної моделі для подальшого аналізу.

У другому експерименті моделювалася ситуація, коли середня складність матеріалу зростає. Очікувалося, що результати продемонструють зниження індивідуальної та особливо загальної ефективності.

У третьому експерименті моделюється ситуація з підвищенням рівня здібностей учнів. Індивідуальна та загальна ефективність дещо підвищуються.

Четвертий експеримент дозволяє оцінити вплив взаємодопомоги учнів. Ми спостерігали очікуване скорочення часу консультування та підвищення ефективності.

Наступна серія експериментів дозволяє оцінити вплив факторів дистанційного навчання.

Можна припустити, що при використанні електронних засобів зв'язку час очікування відповіді викладача скорочується, але збільшується середня кількість запитань студента.

П'ятий експеримент показав адекватність результатів моделювання в порівнянні з даними базового прикладу.

Шостий експеримент дозволив зрозуміти, що при дистанційному навчанні ускладнення матеріалу призводить до значного зниження ефективності навчання та збільшення витрат часу.

Сьомий експеримент показав, що для студентів з високим коефіцієнтом здібностей перехід на дистанційне навчання може призвести до серйозного підвищення ефективності навчання. Крім того, організація легкої взаємодії між студентами в таких групах дає ще більший ефект. Результати восьмого експерименту підтверджують це.

V. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

На початку дослідження було виявлено, що системний аналіз галузі дистанційного навчання повинен включати оцінку сценаріїв взаємодії викладачів між собою, взаємодії студентів та викладачів із технічним персоналом системи. Окрему проблему становить моделювання поведінки менеджерів системи, персоналу, що забезпечує рамки систем дистанційного навчання, правила поведінки інших учасників та адаптацію всієї структури до динамічних змін. Ймовірно, для комп'ютерного моделювання цих процесів необхідно створювати імітаційні моделі, в яких, крім агентного підходу, використовується також дискретно-подійний та системно-динамічний.

VI. ВИСНОВКИ

У даній роботі досліджено проблеми ефективності сучасних систем освіти та існуючі підходи до їх оцінювання. Виявлено, що сучасні умови диктують обов'язкове використання інформаційно-комунікаційних технологій у навчальному процесі студентів. Система освіти є складною і потребує використання засобів моделювання, які дозволять оцінити нелінійну дію різномірних факторів. У даній роботі представлено результати застосування агентного підходу до дослідження проблем оцінки ефективності сучасних систем освіти. Агентна модель дозволяє врахувати основні моделі поведінки учасників взаємодії за допомогою діаграм станів і функцій переходу агентів з одного стану в інший. Опис умов таких переходів дозволяє досліджувати параметри системи, які найбільше впливають на ефективність системи. Представлена агентна модель може стати основою для подальшого дослідження проблем

ефективності систем дистанційної освіти в сфері ІТ за її адекватного динамічного розвитку та з урахуванням інших типів взаємодій у системі.

ЛІТЕРАТУРА

- [1] A. Curaj, L. Deca, L., R.Pricopie “European Higher Education Area: The Impact of Past and Future Policies”. *Springer International Publishing*, 2018, pp.119-141. doi: 10.1007/978-3-319-77407-7
- [2] V.Kupriyanova, T.Estermann, N. Sabic. “European Higher Education Area: The Impact of Past and Future Policies. Efficiency of Universities: Drivers, Enablers and Limitations. Efficiency of Universities: Drivers, Enablers and Limitations”. *Springer International Publishing*, 2018, pp. 603-619. doi:10.1007/978-3-319-77407-7_36
- [3] D.Keegan *Theoretical principles of distance education*. Routledge, 2005.
- [4] R.W.Mendenhall. *A model and principles for effective Internet-based distance education*. Brigham Young University, 2003.
- [5] G.C. Kane, D. Palmer, A. Philips Nguen, D.Kiron, N.Buckley. “Strategy, not technology,drives digital transformation”. *MIT Sloan Manag, Rev.27*, pp.1-27, 2015.
- [6] M. Farajollahi, N. Zarifsanee. “Distance Teaching and Learning in Higher Education: A Conceptual Model,” [Online]. Available: https://www.researchgate.net/publication/221929043_Distance_Teaching_and_Learning_in_Higher_Education_A_Conceptual_Model. Accessed on: November 9, 2022
- [7] A.Borshchev, A.Filippov A. “From system dynamics and discrete event to practical agent based modeling: reasons, techniques, and tools”. *The 22nd International Conference of the System Dynamics Society*, Oxford, England, 2004 [Online]. Available: <https://www.anylogic.com/resources/articles/from-system-dynamics-and-discrete-event-to-practical-agent-based-modeling-reasons-techniques-tools/>. Accessed on: November 9, 2022
- [8] Shanna R Simpson-Singleton, Xiangdong Che. “Agent-Based Modeling And Simulation Approaches In Stem Education Research,” *Journal of International Technology and Information Management*, 2019 Vol. 28. Iss. 3, Article 1. [Online]. Available <https://scholarworks.lib.csusb.edu/jitim/vol28/iss3/1>. Accessed on: November 9, 2022
- [9] M.Puterman, *Markov Decision Processes: Discrete Dynamic Stochastic Programming*. John Wiley. NY, 1994.
- [10] F.Sklar, M. Davies. “Multiagent simulation of learning environments”. *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, 2005.Utrecht, Netherlands. [Online]. Available <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.3351&rep=rep1&type=pdf>. Accessed on: November 9, 2022
- [11] K.Binmore.*Fun and games: A text on game theory*. D. C. Heath and Company, Lexington.MA, USA, 2005

A Practical Approach to Evaluating the Ethics Index of Artificial Intelligence Systems

Oleh Zaritskyi

*Information Technology Security Department
National Aviation University
Kyiv, Ukraine*

Abstract—The article deals with the current issues of classification of the challenges that arise and the principles that should be used, during the development and implementation of artificial intelligence systems. In the subject area of AI ethics, the notion of an AI system ethics index is introduced. The research is purely theoretical in nature, summarizing existing ideas and principles, for the first time putting forward the idea of a quantitative assessment of the question of the ethics of artificial intelligence.

Keywords—*AI System, AI Ethics, Ethics index*

I. INTRODUCTION

The intensive development of information technology in the last decade, especially in the field of artificial intelligence and hardware in the form of neurosynaptic and quantum computers, poses new challenges to society in its harmonious development in terms of moral and ethical issues, as well as information security. Numerous AI programs by the world's leading governments published in the past few years also highlight the urgency of the ethical issues that will arise as these technologies develop. “A Next Generation Artificial Intelligence Development Plan” strategizes China AI development by 2030. By 2025 it states that China will have major breakthroughs in AI theory and AI will become the driving force for industrial upgrading and economic restructuring. And by 2030, China will become the world’s major AI innovation center [1]. In USA DARPA announces “\$2B+ investment plan to overcome limitations on AI technology” [2]. The “AI Next program” begins [3]. The Subcommittee on Information Technology of the U.S. House Committee on Oversight and Government Reform publishes a white paper on AI and its impact on policy [4].

The UK government publishes its “AI Sector Deal” which invests 950M pounds (1.2B USD) to support research / education, and enhance the UK’s data infrastructure [5].

Since 2014-15, public, private companies, educational and research institutions have begun to publish various regulatory documents, materials related to the ethical issues of the development, implementation and application of artificial intelligence systems. The importance of moral issues in information systems and AI is also evidenced by the research on ethical issues highlighted in separate sections of the systematic AI index reports [6-8], which highlight several general principles that unite these documents, among them: confidentiality, accountability, transparency, and explainability.

The very fact that such documents appear shows that society is beginning to pay serious attention to such a difficult issue as ethics and human rights in the field of artificial intelligence. However, criticism should be noted, that has arisen from experts in ethics and human rights in connection with the possible ambiguous or inaccurate use of existing terms in this field. The abstract nature of the introduced principles does not allow us to speak about their adequacy from

the point of view of a correct description of the subject area and, accordingly, about the possibility of their use to control compliance with ethical norms in the field of AI.

II. RELATED WORK

Research [6-8], covering more than a hundred papers produced by various organizations on AI ethics, identifies the 12 most frequently cited challenges to AI ethics (tbl.1):

Table I
Challenges to AI ethics

Ethical Challenges	Definition
Accountability	All stakeholders of AI systems are responsible in the moral implications of their use and misuse
Safety	Throughout their operational lifetime, AI systems should not compromise the physical safety or mental integrity of humans
Human Control	It assumes control by the developer and end-user in the development and use of AI systems, respectively
Reliability, Robustness, and Security	All systems designed and used must be reliable in use, resistant to external influences and meet information security standards
Fairness	The development of AI should refrain from using datasets that contain discriminatory biases
Diversity and Inclusion	Understand and respect the interests of all stakeholders impacted by your AI technology
Sustainability	The AI development must ensure the sustainability of our planet is preserved for future
Transparency	An AI system should be able to explain its decision making process in a clear and understandable manner
Interpretability and Explainability	Developed AI systems should be understandable in terms of their internal content (construction) and easily explainable in terms of their functionality
Multi Stakeholder engagement	Involves multiple independent stakeholders in the development and operation of AI systems
Lawfulness and Compliance	All the stakeholders in design of an AI system must always act in accordance with the law and all relevant regulatory regimes
Data Privacy	Users must have the right to manage their data which is used to train and run AI systems

This list is not exhaustive, but shows general trends in AI. Researches show that fairness, interpretability and explainability, transparency are most mentioned across all documents studied.

Research [9] presents core ethical principles (*i.e., respecting autonomy, avoiding harm & doing good, ensuring justice*) and the instrumental principles that primarily link to them. With about 100 sets of principles published as of today, it is easy to get lost in these separate but similar documents, so the “Dynamics of AI Principles” is tool for keeping track of, and systematize, the bewildering and growing number of AI Principles out there. AI principles had been published by private companies, governmental agencies, international organizations, research centers, and professional organizations.

In [10], much attention had been paid to the legal aspects of the development of AI systems from American legislation system. The report addresses issues such as: privacy, innovation policy, liability (civil), liability (criminal), agency, certification, labor and taxation.

In research report [11] authors mentioned that for millennia, waves of technological change have been perceived as a double-edged sword for the economy and labor market, increasing output and wealth but potentially reducing pay and job opportunities for typical workers. Thus, the study

emphasizes the question how has AI affected socioeconomic relationships? We do not see an unambiguous answer. Perhaps the impact on the economy and the labor market is not as noticeable as expected, because AI has been localized in certain industries and countries and does not have the proper level of implementation, i.e. we did not get a critical mass.

Thus, an analysis of recent research suggests several main areas of concern for ethics and human rights scholars: information security (human rights, adequate historical data for learning samples, etc.) and the impact on human and social well-being.

III. METHODOLOGY

The research methodology involves the study of the key causes of disagreement between the fields of research in artificial intelligence and ethics, as well as the classification (formalization) of the basic concepts of the field of study.

Obviously, all the disagreements between AI specialists and ethicists in its classical sense arise from different interpretations of AI terms in terms of ethics. It is necessary to turn to the definition of ethics and the tasks it addresses in a broad and narrow sense.

Ethics is a philosophical discipline whose subject is morality. Ethics has two main functions – moral-educational and cognitive-educational, so two areas could be distinguished in ethics - normative ethics, aimed at teaching about life, and theoretical ethics, studying morality. [12].

Theoretical ethics is a scientific discipline that examines morality as a special social phenomenon, finds out what it is, how morality differs from other social phenomena. Theoretical ethics studies the origin, historical development, regularities of functioning, social role and other aspects of morality. Its methodological basis is the knowledge, concepts and ideas concerning the scientific knowledge of morality.

Normative ethics searches for a principle (or principles) that governs human behavior, guides one's actions, establishes criteria for evaluating the moral good, and a rule that can act as a general principle for all cases.

Applied (practical) ethics studies particular problems and the application of moral ideas and principles articulated in normative ethics to specific situations of moral choice. Applied ethics interacts closely with the social and political sciences and has a number of sections, e.g. business ethics, medical ethics, computer ethics, etc.

Obviously, AI ethics could be classified as a section of applied (normative) ethics, which is very close to information (computer) ethics. Let us make a brief excursion into information ethics and consider its principles.

Broadly speaking, computer ("information" or "cyber ethics") ethics investigates the behavior of people who use information systems, based on which appropriate moral precepts and norms of behavior are developed [13].

Computer ethics covers almost all spheres of human activity and deals with technical, moral, legal, social, political and philosophical issues. The problems analyzed in it could be roughly divided into several classes:

- 1) problems associated with the development of moral codes for users and developers;
- 2) problems of protection of property rights, copyrights and basic human rights (rights to privacy and freedom of speech, obtaining and using information, the right to work, privacy and personal data, etc.) as applied to the field of information technology;

3) cyber security, determination of the status of incidents and crimes, that is, predominantly legal problems, as a rule, formalized in the form of national legislative acts on information security.

Principles developed in computer ethics:

- 1) privacy – a person's right **to autonomy and freedom in private life**, the right to be protected from intrusion by authorities and others;
- 2) accuracy (accuracy) – compliance with the norms related to the accurate execution of the instructions for the operation of systems and information processing, honest and socially responsible attitude to their duties;
- 3) property – inviolability of private property. Adherence to this principle means observance of the right of ownership of information and copyright norms;
- 4) accessibility – the right of citizens to information, its accessibility at any time and in any place.

The principles of information (computer) ethics developed are very similar to the ethical principles of AI that we reviewed in the literature review, but they relate only to data processing and issues of security and ownership.

The only issue is that not all these principles have been systematized in the framework of a corresponding standard and each developer tries to take into account all possible cases of AI impact on society, which leads to their duplication, different interpretation of sometimes the same principles and challenges. For the same reason various so to speak analytics are also mixed, for example, direct impact on society and incorrect historical data, etc.

The main ethical challenges could be divided into several main groups (fig.1).

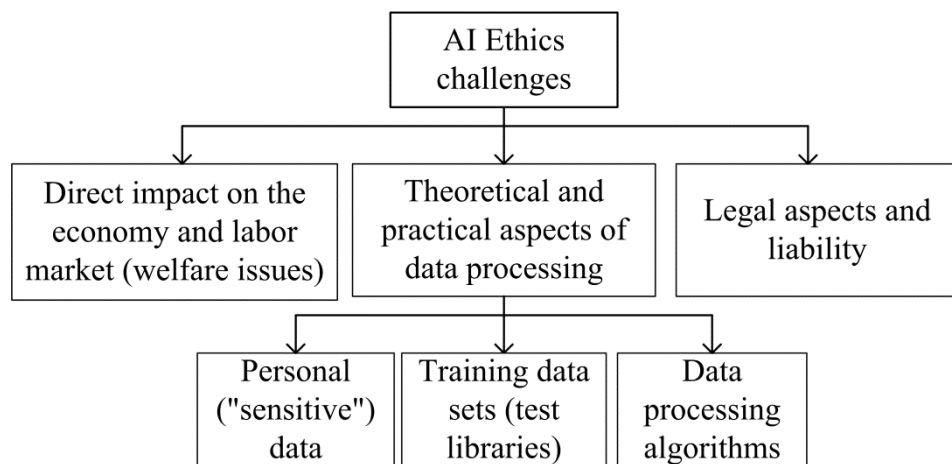


Fig.1. AI ethics challenges

IV. RESULTS

All challenges and problematic issues that arise in one way or another during the development of AI systems can be attributed either to the issues of data and algorithms, or to those that affect the economy and society, or conflict with existing legal norms. Very often the challenges are complex and can simultaneously positively affect the economy as a whole, while also causing social contradictions, such as a couple of productivity and employment issues.

In the aspect of data processing, special attention should be paid to the creation of test (training sets) of data, which require strict adherence to the principles: data neutrality, representative data, accuracy, reliability, openness and diversity.

Among the many approaches to classifying ethical principles in AI issues, the author would distinguish four large global groups (fig.2).

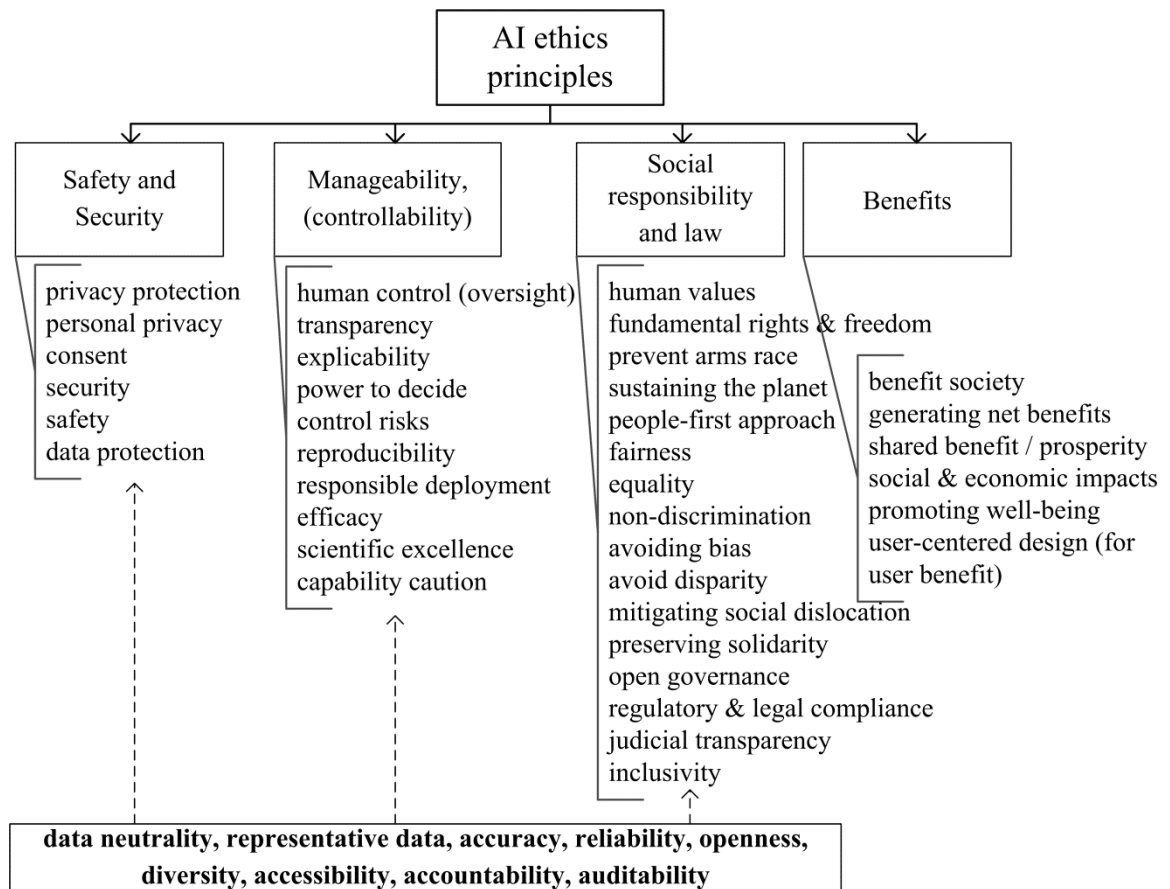


Fig.2. AI ethics principles

Figure 2 shows the classification of ethical principles into four main groups, as well as the relationship of these groups to the principles described in papers [9-11].

Group “Safety and Security” includes all the principles that describe the security and protection of both data and its accuracy and reliability in order to create a secure information technology. A very important issue is the adequacy of historical data to create correct training samples, which affects both the manageability of AI and social responsibility.

Group “Manageability (Controllability)” describes the principles that must be followed to create AI software that is manageable, efficient, understandable to the end user, and controllable by the end user. The principles of this group also imply a cautious attitude toward the prospective capabilities of the AI system are being creating, which have not yet been fully clarified by the developer. The development should be conducted at a high scientific and technical level. The system must be reproducible under different conditions. The developer must take into account all risks in operating conditions.

The group “Social Responsibility and law” includes principles that characterize the AI system to be developed in terms of compliance with social and legal norms of society.

Group “Benefits” describes the principles for assessing the usefulness of the developed AI system from both a tangible and intangible point of view for the user and society as a whole.

There is a group of principles (bottom of Figure 2): data neutrality, representative data, accuracy, reliability, openness, diversity, accessibility, accountability, auditability that are

common to the three groups. These principles relate to data, but their implementation lays the foundation for addressing safety, manageability, and social responsibility.

Using the developed classifiers of challenges and principles, let me introduce an index of ethics of the AI system (1):

$$E_{AI} = \sum_{i=1}^N (\pm P_i + P_i^{IN}); \quad (1)$$

$$\lim_{+P_i \rightarrow \max; -P_i \rightarrow 0} E_{AI} \rightarrow \max$$

$\pm P_i$ - is an evaluation of the principles that the AI system fulfills, (+) means positive impact, (-) – negative;

P_i^{IN} - is a principle (initiative, IN) necessary to compensate for negative influences.

Due to the lack of digitized ethical values in existing approaches, which are more descriptive, it is practically impossible to use historical statistics for quantitative comparison of different approaches.

V. DISCUSSION

The main condition for using formula (1): the principles of development must correlate with the challenges posed by the development and minimize their negative impact on society from an ethical point of view. For example, automation can lead to increased productivity and, as a consequence, to job losses. This is a serious challenge with the highest ethical rating. We can rate the principle – fundamental rights and freedoms (right to work) on a maximum scale, for example – 5 points, $P_i = -5$. At the same time, it will reduce the length of the working day, improving social conditions (if there is such a point in state programs, for example), so we can assess this initiative also on the maximum scale – $P_i^{IN} = +5$. Thus, we are talking about the implementation of initiatives (recommendations for legislatures and businesses), which can minimize the impact of negative factors (challenges). From the point of view of possible limitations in the use of the proposed method, at this stage of the study it is very difficult to determine the factors that can serve as limitations and go beyond the limits of applicability.

VI. FUTURE RESEARCH

The author's top-level classification of AI ethics principles could be detailed by independent research and introduced as a standard after agreement with all stakeholders. The Index of AI Ethics is been considered as a general approach to the evaluation of developed AI systems. It requires further study in terms of a detailed classification of principles in the form of a final list, which will be included in the upper level groups proposed by the author. The development of evaluation scales for quantitative assessment of the ethics of AI systems is also been envisaged. The results of research in the field of ethics of artificial intelligence are closely intertwined with research and approaches to quantify the technological singularity proposed in work [14].

VII. CONCLUSION

AI ethics in a broad sense could be seen as an applied part of general ethics, which examines the behavior of people who develop and use AI systems, as well as the impact of these systems on society. The result of this study is the principles and norms of morality designed both to solve private practical problems in the development process and to realize the harmonious development of society with maximum ethical benefit.

VIII. REFERENCES

- [1] A Next Generation Artificial Intelligence Development Plan, China copyright and media, Aug. 2017. [Online]. Available: <https://chinacopyrightandmedia.wordpress.com/2017/07/20/a-next-generation-artificial-intelligence-development-plan/>.
- [2] DARPA announces “\$2B+ investment plan to overcome limitations on AI technology”, Defense Advanced Research Projects Agency, 2018. [Online]. Available: <https://www.darpa.mil/news-events/2018-09-07>.
- [3] AI Next program begins, Defense Advanced Research Projects Agency, 2018. [Online]. Available: <https://www.darpa.mil/work-with-us/ai-next-campaign>.
- [4] Rise of the Machines: Artificial Intelligence and its Growing Impact on U.S. Policy, United states. Congress. House. Committee on oversight and government reform, 2007. [Online]. Available: <https://www.hsdl.org/?abstract&did=816362>
- [5] AI Sector Deal. Policy pape. Gov.uk.
<https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal>.
- [6] Y. Shoham, R. Perrault, E. Brynjolfsson, J. Clark, J. Manyika, J.C. Niebles, T. Lyons, J. Etchemendy, B. Grosz and Z. Bauer. Artificial intelligence Index. 2018 Annual Report. Steering Committee. Stanford University. 2018.
- [7] Y. Shoham, E. Brynjolfsson, J. Clark, J. Manyika, J.C. Niebles, T. Lyons, J. Etchemendy, B. Grosz and S. Mishra. Artificial intelligence Index. 2019 Annual Report. Steering Committee. Stanford University. 2019.
- [8] Y. Shoham, E. Brynjolfsson, J. Clark, J. Manyika, J.C. Niebles, T. Lyons, J. Etchemendy, B. Grosz and S. Mishra. Artificial intelligence Index. 2021 Annual Report. Steering Committee. Stanford University. 2021.
- [9] TOOLBOX: Dynamics of AI Principles, AI ETHICS LAB. Aiethicslab.com.
<https://aiethicslab.com/big-picture/>.
- [10] Artificial intelligence and life in 2030. One hundred year study on artificial intelligence. Report of the 2015. Study panel, Stanford University, 2016. [Online]. Available: <https://ai100.stanford.edu/2016-report>.
- [11] Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report. Stanford University, 2021. [Online]. Available: <https://ai100.stanford.edu/2021-report/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence>.
- [12] А.В. Разин. Этика: Учебник для вузов / 3-е изд., перераб. — М.: Академический проект, 2006.
- [13] T. Bynum. Computer and Information Ethics. The Stanford Encyclopedia of Philosophy (Spring 2011 Edition), 2018.
- [14] O. Zaritskyi, O. Ponomarenko. Quantitative assessment of technological singularity, The International Scientific and Technical Journal Problems of control and informatics, 2022. - №1. - P.93-111.
DOI: <http://doi.org/10.34229/1028-0979-2022-1-9>.

Розподілені системи виявлення зловмисного програмного забезпечення

Богдан Савенко

*Кафедра комп'ютерної інженерії та інформаційних систем
Хмельницький національний університет
Хмельницький, Україна*

Анотація—В роботі пропонується використовувати різноманітні антивірусні засоби та системи виявлення вторгнень на різних етапах та рівнях проходження в комп'ютерних системах і мережах зловмисного програмного забезпечення. Розроблено розподілену систему для виявлення зловмисного програмного забезпечення, в якій синтезовано можливість до самоорганізації та зміни місця знаходження центру прийняття рішень системи. Експериментальні дослідження підтвердили можливість реалізації запропонованих рішень. Отримана розподілена система має достатній рівень стійкості для забезпечення її функціонування і виконання імплементованих в неї методів виявлення зловмисного програмного забезпечення.

Ключові слова—*Розподілені системи, Зловмисне програмне забезпечення, Системи виявлення вторгнень, Антивірусні програмні засоби.*

I. ВСТУП

Розробка та активне використання і поширення зловмисного програмного забезпечення (ЗПЗ), крім створення проблем користувачам щодо вирішення задач в комп'ютерних системах та мережах (КСМ), також блокує доступ до КСМ, виводить їх з ладу на певний час та для усунення наслідків вимагає витрат часу кваліфікованого персоналу. Використання сучасних антивірусних засобів (АВЗ) та систем виявлення вторгнень (СВВ) є досить ефективним, але використовувані при створенні ЗПЗ нові методи дають змогу зловмисникам частково обходити такі засоби. Крім того, робота АВЗ та СВВ потребує залучення користувача до процесу адміністрування подій, про які інформуватимуть ці засоби і системи у випадку виявлення ЗПЗ чи підозри на наявність ЗПЗ. А це призводитиме до витрат часу, особливо коли користувач не має відповідної кваліфікації. Практики із забезпечення безпеки та захисту КСМ відповідними АВЗ та СВВ підтверджують необхідність застосування комплексних систем, які б включали багаторівневість, багатоетапність та різноманітність. Різноманітність АВЗ та СВВ, відомості про які невідомі зловмиснику, суттєво покращують безпеку та захист КСМ. Також, важливим для таких засобів є можливість до перебудови своєї архітектури в залежності від стану безпеки в КСМ. Тому, перспективним напрямом для дослідження є створення розподілених систем виявлення ЗПЗ в КСМ, архітектура яких надавала б змогу до самоорганізації такої системи щодо подальших кроків при виявленні ЗПЗ чи підозри на наявність ЗПЗ та зміни місця знаходження центру системи для уникнення атак на нього.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Зловмисне програмне забезпечення, незважаючи на різноманітність засобів протидії, продовжує розвиватись, про що підтверджують результати незалежних дослідницьких лабораторій [1, 2]. Методи виявлення ЗПЗ зорієнтовані переважно на класифікацію об'єктів, які надходять в систему. Наприклад, метод виявлення, поданий в роботі [3] має на меті врахування поведінки і виділення в ній зловмисних дій для подальшої класифікації. Також, досліджуються кібератаки з використанням різних відомих математичних методів, зокрема в 2022 International Conference on Innovative Solutions in Software Engineering Ivano-Frankivsk, Ukraine, November 29-30, 2022

роботі [4] подано застосування еволюційних алгоритмів для їх виявлення. Крім того, в роботі [5] подано метод виявлення ЗПЗ з використанням дослідження команд прикладного програмного інтерфейсу. Такі методи активно розробляються та вдосконалюються дослідниками. Перспективними методами є методи з використанням штучних нейронних мереж, але точність виявлення при їх застосуванні потребує дослідження. Метод з цієї групи методів подано в роботі [6]. Таким чином, методи виявлення ЗПЗ є досить різноманітними і охоплюють різні спрямування, а також використовують в своїх етапах чи кроках відомі математичні методи. При цьому, подаючи розроблені методи та досліджені ефективності цих методів, дослідники переважно не вказують на засоби, в які вони імплементують розроблені методи. Узгодження функціонування систем з імплементованими в неї методами виявлення ЗПЗ в залежності від станів безпеки КСМ, в яких вони функціонують впливатиме на результат виявлення. Тому, перспективним напрямком для дослідження є створення розподілених систем з архітектурою, в якій буде синтезована самоорганізація та можливість зміни місця знаходження центру системи для уникнення його виявлення. А, також, узгодження поведінки такої системи щодо подальших її кроків з результатами роботи імплементованих в неї методів виявлення ЗПЗ.

III. МОДЕЛІ АРХІТЕКТУРИ РОЗПОДІЛЕНОЇ СИСТЕМИ

Засоби систем виявлення зловмисного програмного забезпечення в комп'ютерних мережах, а також в їх хостах, повинні базуватись не тільки на сучасних актуальних методах виявлення, але і бути імплементованими в такі архітектури засобів, які б залучали свої елементи до покращення виявлення сумісно із методами виявлення. Реагування систем виявлення ЗПЗ в комп'ютерних мережах завдяки динамічній перебудові своїх архітектур в умовах зловмисних впливів та аномальних проявів створює додаткові перешкоди для зловмисників та ЗПЗ. Така динамічна перебудова архітектури повинна координуватись та узгоджуватись із застосуванням методів виявлення. Створення для зловмисників та ЗПЗ перешкод в розумінні функціонування та поведінки засобів виявлення на архітектурному рівні надає перевагу користувачам КСМ. Досягнення переваги потребує крім методів, які орієнтовані безпосередньо на виявлення ЗПЗ, забезпечити в складі засобів виявлення компоненти або елементи, які змінюватимуть архітектуру системи виявлення сумісно з методами виявлення, але при цьому вони не орієнтовані саме на виявлення конкретних типів ЗПЗ чи конкретного ЗПЗ. Такі компоненти чи елементи повинні забезпечувати функціонування системи виявлення ЗПЗ без втручання адміністратора системи чи користувача при прийнятті рішень щодо подальшого функціонування в умовах впливів ЗПЗ і не бути прогнозованими в своїх подальших діях для зловмисників та користувачів. Модель архітектури M_S частково централізованої системи S задамо згідно її компонентів та зв'язків між ними так:

$$M_S = \langle S, G_S \rangle, \quad (1)$$

де G_S – граф, що відображає зв'язки між компонентами частково централізованої системи S .

Враховуючи поділ компонент системи на дві підмножини за критерієм наявності центру в них і без нього, отримуємо уточнену модель архітектури $M_{S,k}$ частково централізованої системи S згідно формули:

$$M_{S,k} = \langle \langle S_1, S_2, \dots, S_k \rangle, \langle S_{k+1}, S_{k+2}, \dots, S_N \rangle, G_S \rangle, \quad (2)$$

де G_S – граф, що відображає зв'язки між компонентами частково централізованої системи S ;
 k – кількість компонент системи, в яких може бути центр системи;

$S = S_1 \cup S_2 \cup \dots \cup S_k \cup \dots \cup S_N$; N – кількість компонент в системі, які встановлені в комп'ютерні станції в мережі.

В запропонованій таким чином моделі архітектури $M_{S,k}$ частково централізованої системи, крім розподілу компонент системи, які задано підмножинами і відповідно вершинами, в залежності від можливостей містити центр, виділено також три типи зв'язків між компонентами, які поєднують компоненти з центром системи, компоненти без центру системи та компоненти з центром і без центру між собою. Граф G_S при такому заданні є повним, тобто з'єднання між компонентами система наявні між ними усіма. Але для ефективнішої роботи та приховування можливостей системи з'єднання між компонентами системи можуть бути задані різними деревами графа G_S і, таким чином, їх кількість зменшуватиметься, а також, приховуватимуться від зловмисника або ЗПЗ очікувані повідомлення. Визначення варіантів дерев графа G_S встановлюватиметься центром прийняття рішень системи.

Задамо варіанти з'єднань між різного типу компонентами системи через відповідні їм функції. Таких функцій буде три і будуть, також, відповідні їм обернені. Визначення функцій необхідно для подальшої деталізації моделі архітектури системи та розроблення відповідного протоколу комунікації між ними. Функції задамо формулами:

$$\begin{aligned} F_1: S_i \xrightarrow{G_S} S_j, 0 < i, j \leq k, i \neq j; \\ F_2: S_i \xrightarrow{G_S} S_j, 0 < i \leq k, k < j \leq N; \\ F_3: S_i \xrightarrow{G_S} S_j, k < i, j \leq N, i \neq j, \end{aligned} \quad (3)$$

де $S_i - i$ – та компонента системи; j – та компонента системи; k – кількість компонент системи, в яких може бути центр системи; $S = S_1 \cup S_2 \cup \dots \cup S_k \cup \dots \cup S_N$; N – кількість компонент в системі, які встановлені в комп'ютерні станції в мережі; G_S – граф, що відображає зв'язки між компонентами частково централізованої системи S .

Архітектура частково централізованої системи S задана її моделлю архітектури відповідає вимогам щодо можливості динамічної зміни конфігурації, поділу центру прийняття рішень, розподілу компонентів за можливостями з наявності центру прийняття рішень в них і, тому, є основою для подальшого синтезу в ній властивостей адаптивності і самоорганізованості, реалізація яких буде здійснюватись безпосередньо в компонентах системи, в основному тих з них, в яких знаходиться центр прийняття рішень системи.

IV. РЕЗУЛЬТАТИ

Розроблена архітектура частково централізованої системи S була реалізована у вигляді програмного забезпечення, яке поєднало комп'ютерні станції в комп'ютерній мережі. Критеріями оцінювання якості розробленої архітектури були такі: час, який витрачено на комунікацію між компонентами системи; час, витрачений на визначення подальших кроків системи, тобто на прийняття рішень щодо подальших кроків; час, який витрачено на забезпечення міграції центру прийняття рішень; коефіцієнт деградації системи; коефіцієнт деградації компонентів та елементів системи. Отримані результати загального критерію щодо можливості функціонування системи, як самоорганізованої системи, становлять більше 85%, що є прийнятним для завдань, які плануються нею для виконання. Безпосередньо в розроблену систему було імплементовано метод виявлення метаморфних вірусів. Метрикою визначення якості виявлення метаморфних вірусів засобами частково централізованої системи та розробленого методу було використано бінарну класифікацію згідно помилок першого та другого роду, а також час, який було витрачено на їх виявлення в різних

комп'ютерних станціях в мережі. Результати з виявлення однакових метаморфних вірусів, які перебували в різних комп'ютерних станціях в мережі відрізняються на 14% за часом, що є прийнятним, враховуючи час обробки подій в частково централізованій системі.

V. ОБГОВОРЕННЯ

Отримані результати дали змогу підтвердити можливість практичної реалізації розроблених та запропонованих рішень. Згідно отриманих рішень є можливим реалізація розподілених систем з імплементованими в них методами виявлення ЗПЗ та створення на їх основі АВЗ. Розроблена система має обмеження щодо використання. Вона може застосуватись сумісно з іншими СВВ та АВЗ. Місцем її застосування є комп'ютерні станції в мережі. Її засоби виявлення спрямовані на здійснення аналізу щодо трафіку в комп'ютерній мережі, виявлення метаморфних вірусів та відслідковування зловмисних дій в комп'ютерних станціях, а також порівняння проявів зловмисних дій у вузлах в мережі.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Напрямами подальших досліджень є удосконалення та деталізація архітектури системи, наповнення її новими методами виявлення ЗПЗ, а також оцінювання якості отриманих рішень з використанням відповідних критеріїв та метрик.

VII. ВИСНОВКИ

У результаті проведеного дослідження було розроблено частково централізовану систему виявлення ЗПЗ, в архітектурі якої синтезовано принцип самоорганізації. А також, в неї імплементовано метод виявлення метаморфних вірусів. Функціонування системи та виконання роботи імплементованого методу виявлення метаморфних вірусів узгоджуються в процесі функціонування в залежності від рівня безпеки в КСМ.

ЛІТЕРАТУРА

1. AVTest (May 2022). Malware Statistics [Online]. Available: <https://www.av-test.org/en/statistics/malware/>
2. ITU (2022). ITU Statistics [Online]. Available: <http://www.itu.int/ict/statistics>
3. M. Geden and J. Nappa, "Classification of malware families based on runtime behaviour," *International Symposium on Cyberspace Safety and Security*, 2018, pp. 33-48
4. S. Lysenko, K. Bobrovnikova, R. Shchuka and O. Savenko, "A Cyberattacks Detection Technique Based on Evolutionary Algorithms," *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 2020, pp. 127-132, doi: 10.1109/DESSERT50317.2020.9125016
5. O. Savenko, A. Nicheporuk, I. Hurman and S. Lysenko, "Dynamic Signature-based Malware Detection Technique Based on API Call Tracing," 2019, CEUR-WS, vol. 2393, pp. 633-643
6. H. Zhou "Malware Detection with Neural Network Using Combined Features," *Communications in Computer and Information Science*, Springer, 2019, vol 970, doi:https://doi.org/10.1007/978-981-13-6621-5_8

Метод керування та самоорганізована система планування маршрутів БПЛА

Олександр Мельниченко

*Кафедра комп'ютерної інженерії та інформаційних систем
Хмельницький національний університет
Хмельницький, Україна*

Анотація—Розвиток технологій із використанням БПЛА (безпілотних літальних апаратів), потребує стандартизації методів управління як для наземних комплексів, так і для літальних апаратів та платформ. Збільшення кількості літальних апаратів потребує пошуку методів автономної навігації, способів утворення груп та мереж. Тому, питання впровадження нових методів керування є актуальним. Експериментальні дослідження підтвердили можливість реалізації запропонованих рішень. Отримана самоорганізована система має достатній рівень стійкості для забезпечення її функціонування і виконання впроваджених в неї методів управління.

Ключові слова— *Технології самонавчання; Безпілотні літальні апарати; Самоорганізована система; Методи керування; Планування маршрутів.*

I. ВСТУП

Упродовж останніх десятиліть, у всьому світі активно розробляються безпілотні авіаційні комплекси (БПЛА) як системоутворюючі елементи БПЛА. Водночас пріоритет розроблення залишається за інформаційним програмними комплексами, завданням яких є управління та моніторинг БПЛА. Однією зі галузей застосування БПЛА є сільське господарство. Інтелектуальна система моніторингу сільського господарства або просто інтелектуальне сільське господарство є новою технологічною концепцією, у якій дані з декількох сільськогосподарських областей, від дрібного до великого масштабу, отримуються з використанням БПЛА. Зібрані дані аналізуються експертами, щоб робити короткострокові та довгострокові висновки про погодні умови, родючість ґрунтів, поточну якість сільськогосподарських культур, кількість плодів тощо. Проте наявні технології управління погано справляються із завданням обчислення кількості плодів у фермерських господарствах. Для подібних рішень потрібен інтелектуальний програмний комплекс, який здатен будувати маршрути на основі технології самонавчання та об'єднувати декілька БПЛА в групу. Тому, перспективним напрямком для реалізації завдання обчислення плодів є розроблення методів керування і самоорганізованої системи планування маршрутів для декількох БПЛА.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Бурхливий розвиток інформаційних технологій упродовж останніх років призвів до появи низки стартапів, що пропонують різноманітні інформаційні системи в галузі «розумного» сільського господарства. Одним із ключових чинників успішності впровадження таких систем є забезпечення максимальної автоматизації завдань системи «розумного» сільського господарства загалом [1-4]. Перспективним підходом до реалізації подібної інформаційної системи є використання штучних нейронних мереж у засобах управління декількома БПЛА [2-5]. Проте, такі підходи є досить різноманітними, важкими в реалізації на малих обчислювальних пристроях та дають можливості автоматично побудувати вигідні з точки зору відстані та ресурсозатратності маршрути [6, 7]. Заразом подібні засоби та технології управління зазвичай не враховують погодні умови робочого середовища та цілісність

отриманих даних. Тому, з огляду на вище вказані проблеми, актуальною науковою задачею є створення автономних систем з архітектурою, у якій буде синтезована самоорганізація та можливість інтегрувати системи сторонніх розробників для покращення результатів. Для розв'язку цієї задачі в даній роботі пропонується новий метод для здійснення обльоту групою БПЛА фруктового саду та обчислення кількості плодів на деревах в автоматичному режимі.

III. МЕТОД КЕРУВАННЯ І САМООРГАНІЗОВАНА СИСТЕМА ПЛАНУВАННЯ МАРШРУТІВ

На основі проведених експериментів в реальному сільськогосподарському середовищі встановлено, що для обчислення кількості плодів з одного дерева доцільно об'єднати декілька БПЛА одну в групу. Як наслідок, отримано певну кількість дронів зі своїми координатами у просторі. На мапу місцевості, що подається координатною сіткою, було нанесено початкові координати для обльоту одного ряду робочого середовища. В результаті отримано певну матрицю станів, що відображає фізичні пристрої з їхніми координатами та точками виконаних завдань. Таке подання компонентів через стани, в яких можуть перебувати апаратні пристрої під час функціонування, дає змогу визначати стан роботи та проводити моніторинг критичних ситуацій чи збоїв. Такі характеристики дали змогу сформувати множину координат для побудови квазіоптимального маршруту. Аналіз інформації з матриці станів дає можливість інформаційній системі приймати рішення про подальші дії кожного пристрою. Для управління групою БПЛА розроблено централізований програмний модуль. Важливим елементом модуля є його можливість бути відкритим для інтеграції інших підсистем. Об'єкт дослідження у сільському господарстві зумовлений багатьма динамічними зовнішніми чинниками, а саме різними погодними умовами та обмеженнями ресурсів потужності та пропускної здатності БПЛА. Такі чинники несуть негативний вплив на якість розпізнавання плодів, і як наслідок, на правильність їхнього підрахунку. А це в свою чергу зменшує час виконання роботи, зменшує відстань, яку БПЛА може облетіти, зменшує обчислювальні можливості та обмежує комунікації з програмною системою. Водночас для уникнення дублювання плодів та розпізнавання сторонніх об'єктів БПЛА повинні виконувати обліт економічно ефективними маршрутами. Однак засіб планування маршрутів має використовувати ефективні підходи до планування шляху, які будуть спрямовані на мінімізацію загальної довжини польоту.

Отже, для досягнення мети дослідження та з огляду на вище вказані обмеження було розроблено метод планування маршрутів з технологією самонавчання. Даний метод розроблено ґрунтуючись на алгоритмі Q-Learning (QL) підходу штучного інтелекту Reinforcement Learning (RL). Метод планування маршрутів для управління групою БПЛА реалізовано у вигляді централізованого програмного модуля. Модуль, що реалізує метод планування маршрутів, генерує матрицю станів з координатами групи БПЛА, що на основі використання алгоритму QL, дає змогу усім пристроям в тривимірному просторі корегувати власну поведінку через взаємодію з робочим середовищем. Підсистема не використовує будь-яких моделей машинного навчання тільки значення матриці станів і працює відповідно до підходу "дія-винагорода". Всі апаратні пристрої в групі виконують свою роботу допоки не досягнуть "позитивного завершення" роботи самоорганізованої системи. Важливим елементом модуля є його можливість бути відкритим для інтеграції інших підсистем. Приклад результату виконання методу подано в табл. 1.

Таблиця 1

Приклад результату автоматичного корегування маршрутів

Стан	Дія 1	Дія 2	...	Дія n
------	-------	-------	-----	-------

S_1	$Q(S_1, A_1)$	$Q(S_1, A_2)$...	$Q(S_1, A_n)$
S_2	$Q(S_2, A_1)$	$Q(S_2, A_2)$...	$Q(S_2, A_n)$
... (S_n, A_n)
S_n	$Q(S_n, A_1)$	$Q(S_n, A_2)$...	

Процес автоматичного створення та корегування маршруту групи БПЛА відбувається ітеративно. З огляду на ітеративні дії кожного дрону, Q-значення в таблиці постійно оновлюються, допоки не буде виконано критерій вигідності (найменша відстань) автоматично побудованого маршруту. Відповідно, метрикою оцінювання якості планування маршруту групи БПЛА є відстань, яку проходить група за новобудованим маршрутом за одну робочу місію в робочому середовищі. Як наслідок, новобудований маршрут вважатиметься квазіоптимальним, оскільки умови різних робочих середовищ є різними.

IV. РЕЗУЛЬТАТИ

В рамках дослідження проведено оцінювання запропонованого методу; зокрема, здійснено порівняння результатів побудови маршруту запропонованим методом з аналогами у відокремлених зонах фруктових садів.

Таблиця 2

Порівняння відстаней побудованих маршрутів різними методами у відокремлених зонах робочого середовища, м

Стан	Метод 1 [1]	Метод 2 [2]	Запропонований метод
Зона робочого середовища 1	8,95	9,28	9,15
Зона 2	8,43	8,14	7,95
Зона 3	6,73	6,7	6,71
Зона 4	8,4	8,25	8,35
Зона 5	11,6	11,45	11,29
Зона 6	10,76	10,8	10,58
Зона 7	16,1	15,75	15,85
Робоче середовище загалом	70,97	70,37	69,88

Відповідно до табл. 2, для трьох із десяти зон робочого середовища запропонований метод дав змогу побудувати підмаршрути меншої відстані, що виражена у метрах. Відстань побудованого маршруту запропонованим методом у робочому середовищі загалом становила 69,88 м, що на 1,54% та 0,7% менше за відстань, отриманої за Методом 1 [1] та Методом 2 [2], відповідно. З огляду на отримані результати (табл. 2), запропонований метод корегування та самоорганізована система планування маршрутів БПЛА дали змогу зменшити відстань маршруту групи БПЛА, і як наслідок, забезпечили підвищення його якості, що свідчить про досягнення мети роботи.

V. ОБГОВОРЕННЯ

Запропонована самоорганізована система планування маршрутів БПЛА дає змогу будувати нові та корегувати чинні маршрути через ітеративне навчання та взаємодію з робочим середовищем. Реалізація самоорганізованої системи забезпечує розв'язання таких завдань: 1) зменшення часу затримки протягом виконання завдання; 2) зменшення відстані під час виконання переміщення у робочому середовищі; 3) збільшення кількості даних, яку БПЛА може опрацювати під час роботи у режимі реального часу. Водночас застосування алгоритму Q-learning до побудови нових маршрутів дає змогу отримати квазіоптимальні значення відстаней між точками переміщення апаратних пристроїв у робочому середовищі, що є критерієм якості побудованих маршрутів. Квазіоптимальність маршрутів, які буде самоорганізована система, є її поточним обмеженням, усунення якого слугуватиме об'єктом подальших досліджень.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Напрямами подальших досліджень є удосконалення системи самонавчання та корегування маршрутів групи БПЛА в робочому середовищі для підвищення якості побудованих маршрутів.

VII. ВИСНОВКИ

Запропоновані метод корегування та самоорганізована система планування маршрутів БПЛА загалом дають змогу будувати нові та корегувати чинні маршрути через ітеративне навчання та взаємодію з робочим середовищем. Реалізація самоорганізованої системи забезпечує розв'язання таких завдань: 1) зменшення часу затримки протягом виконання завдання; 2) зменшення відстані під час виконання переміщення у робочому середовищі; 3) збільшення кількості даних, яку БПЛА може опрацювати під час роботи у режимі реального часу. З огляду на отримані результати (табл. 2), запропонований метод корегування та самоорганізована система планування маршрутів БПЛА дали змогу зменшити відстань маршруту групи БПЛА, і як наслідок, забезпечили підвищення його якості, що свідчить про досягнення мети роботи.

ЛІТЕРАТУРА

1. J. Dong, K. Ota and M. Dong, "Real-Time Survivor Detection in UAV Thermal Imagery Based on Deep Learning," *2020 16th International Conference on Mobility, Sensing and Networking (MSN)*, 2020, pp. 352-359, doi: 10.1109/MSN50589.2020.00065.
2. U. F. Ukaegbu, L. K. Tartibu, M. O. Okwu, I. O. Olayode, "Development of a Light-Weight Unmanned Aerial Vehicle for Precision Agriculture," *Sensors*, 2021, vol 21, no. 13: 4417, doi: <https://doi.org/10.3390/s21134417>
3. W. Zhao, W. Yamada, T. Li, M. Digman, T. Runge "Augmenting Crop Detection for Precision Agriculture with Deep Visual Transfer Learning – A Case Study of Bale Detection," *Remote Sensing*, 2021, vol 13(1), doi:23: <https://doi.org/10.3390/rs13010023>
4. Foodtech startups and venture capital. Five Seasons Ventures, dealroom.co, Q2 2022 [Online]. Available: <https://dealroom.co/uploaded/2022/07/Foodtech-Q2-2022-report-1.pdf?x39545>
5. Enterprise league. (2022). 19 top agriculture startups that are revolutionizing the industry [Online]. Available: <https://enterpriseleague.com/blog/agriculture-startups/>

6. J. Kim, S. Kim, C. Ju and H. I. Son, "Unmanned Aerial Vehicles in Agriculture: A Review of Perspective of Platform, Control, and Applications," in *IEEE Access*, vol. 7, pp. 105100-105115, 2019, doi: 10.1109/ACCESS.2019.2932119.
7. S. Sharma, S. Solanki, K. Aswal, E. Thakur and I. Malhotra, "Review On Application Of Drone Systems In Agriculture," *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, 2021, pp. 40-45, doi: 10.1109/ISPCC53510.2021.9609383.

Розробка веб-додатку для дистанційного навчання

Ольга Дмитрієва^{1,2}, Ілля Бутенко¹

¹*Кафедра прикладної математики та інформатики
Донецький національний технічний університет
Покровськ, Луцьк, Україна,*

²*Дослідницький центр моделюючих технологій (SimTech)
університету Штутгарта, Штуттгарт, Німеччина*

Анотація—Дослідження присвячене проєктуванню процесів розробки веб-додатку для впровадження дистанційного навчання, реалізації обміну завданнями та освітніми матеріалами між користувачами системи з метою підвищення рівня засвоєння знань.

Ключові слова—*веб-додаток, дистанційне навчання, модель даних, макет, Python, JavaScript, інтерфейс, тестування.*

I. ВСТУП

Нестабільна геополітична ситуація та епідеміологічні обставини останнім часом суттєво вплинули на зростання здобувачів освіти будь якого рівня, які не спроможні отримувати освітній ступінь очно. Актуальною стала проблема здобуття знань на віддаленій дистанції з відсутністю фактичної взаємодії між студентами та викладачами. Як результат - низька мотивованість щодо самостійної систематичної роботи над навчальним матеріалом. Виходом з цієї ситуації стало дистанційне навчання [1], до основних переваг якого можна віднести організацію самостійної роботи в зручному місці, темпі та часі. Крім того, для ефективного навчання студенти не лише повинні мати навички роботи з інструментами сучасних інформаційних систем, але й доступ до спеціалізованих систем з інформаційними навігаторами.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Останні наукові дослідження, проведені на великих вибіркових сукупностях [2], показали відсутність методичної бази для самонавчання, низький рівень сили волі і неорганізованість планування самостійної роботи. Деякі з дослідників [3-4] вважають і підтверджують це проведенням глибоким статистичним аналізом, що на мотивацію студентів при застосуванні дистанційного навчання впливають різні демографічні фактори [3], серед яких поточний рік навчання, попередні досягнення студентів, знайомство з навчальним середовищем і, навіть, стать. В [4-5] до результатів онлайн-опитування було застосовано підходи кореляційного аналізу, t-критерію, однофакторного дисперсійного аналізу, множинного лінійного регресійного аналізу та моделювання структурних рівнянь. На підставі проведених досліджень було доведено, що такі психофізичні фактори, як стрес, тривога, депресія та нетерпимість до невизначеності негативно корелюють з мотивацією до дистанційного навчання та частотою відвідування онлайн занять. Також за результатами дослідження, наведеного в [4], доведено, що використання синхронного навчання підвищує відвідуваність занять, в той же час асинхронне дистанційне навчання підвищує мотивацію. І більшість студентів у якості основних недоліків віддаленого навчання [6-7] вказують на обмеженість особистого спілкування. Виходячи з проведеного аналізу останніх досліджень і публікацій, а також, враховуючи поточну ситуацію в країні, яка обумовлена не лише

карантинними обмеженнями, було прийнято рішення розробити веб-додаток для дистанційного навчання, що забезпечує можливості як синхронного, так і асинхронного відвідування занять, створення особистих і групових чатів для спілкування, автентифікацію користувачів та загальноприйняті положення безпеки.

III. МЕТОДИ

Методи, які застосовувалися при розробці веб-додатку для дистанційної освіти при реалізації процесів проведення дистанційного навчання, спираються на основні положення теорії алгоритмів, структур та баз даних, проектування архітектури та програмного забезпечення. Взаємодію між користувачами в мережі було організовано на підставі основних підходів програмного забезпечення мережевих технологій та протоколів передачі інформації, технологій розробки та супроводження інтернет-сайтів. Загальноприйняті положення безпеки програм та даних, взаємодії з користувачем, його автентифікації було покладено в основу підсистеми захисту інформації.

IV. РЕЗУЛЬТАТИ

Розроблений веб-додаток орієнтований на впровадження дистанційного навчання, реалізацію обміну завданнями та освітніми матеріалами між користувачами системи з метою підвищення засвоєння знань здобувачами освіти. Розроблений продукт відповідає всім вимогам, які висуваються до відповідних застосунків, а саме: забезпечує гнучкість, модульність, паралельність, віддаленість, асинхронність функціонування і соціальну спрямованість. Розробка програмного додатку здійснювалась за допомогою теоретичних положень і практичного досвіду у дисциплінах розробки веб-сервісів. Практичне значення розробки полягає у створенні додатку для можливості впровадження дистанційного навчання з подальшою підтримкою та розширенням функціональності.

V. ОБГОВОРЕННЯ

За допомогою створеного додатку для учасників навчання (студентів) навчальний процес у дистанційному режимі забезпечується в зручному місці, темпі і часі. Студентам надається можливість, в залежності від моделі навчання, спілкуватися як з наставниками, так і один з одним за підтримки сучасних інформаційних технологій.

В якості наставників в додатку передбачені ролі як особистого наставника так і групового. Останній стежить за формуванням та виконанням індивідуальних навчальних планів студентів, надає інформацію щодо вибіркового дисциплін, розробляє розклад занять, організовує консультації, забезпечує доступ до навчально-методичних матеріалів. Особистий наставник забезпечує проведення групових занять та індивідуальних консультацій з відповідної дисципліни, здійснює контроль засвоєності студентами інформації навчального матеріалу та оцінює їх знання. Оскільки за умов дистанційного навчання зростає ймовірність фальсифікації оцінюваних показників навчально-виховного процесу, наявна проблема контролю навчального процесу на дистанційному навчанні. Для вирішення такої проблеми в додатку застосовано спеціальні заходи безпеки. Так, під час дистанційного тестування проводиться додаткова фото ідентифікація, передбачено спеціальні методи перевірки незалежності виконання завдання.

Технічна складова проєкту передбачала розробку моделей даних та макетів. Серед моделей даних окрім загальноприйнятних (навчальний заклад, група, користувач і т.ін.) передбачено моделі завдань, бібліотеки завдань, файлів відповідей, адміністраторів завдань. На основі розроблених моделей даних забезпечено функціонування вбудованого механізму роботи з базами даних так їх таблицями. Розроблені макети для веб-додатку передбачають наявність головної (титульної) сторінки, сторінок реєстрації, логіну, відновлення паролю, налаштування облікового запису, повного списку завдань, створення і відображення

завдання. Програмна реалізація здійснювалася засобами Python [8] з фреймворком для розробки веб-додатків Django, та JavaScript.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Можливі розширення функціональності веб-додатку, що пропонується, шляхом розробки та впровадження додаткового функціоналу, а саме: організації чату між користувачами, проведенням відео конференцій, створенням оголошень, взаємодією з календарем, автоматизованими підходами до оцінювання. Для такого розширення функціональності необхідно розробити додаткові моделі користувачів і даних, описати їх взаємодію між собою, визначити макети сторінок та модифікувати архітектуру додатку.

VII. ВИСНОВКИ

Проведено аналіз предметної області, останніх досліджень та публікацій за обраною темою, розібрано конкурентні рішення, сформовано основні вимоги до характеристик створеного проєкту. Проектування дистанційної системи навчання здійснювалося за допомогою UML-діаграм (класів, станів, прецедентів, послідовностей). Спроектовано моделі даних та макети для використання у розроблюваному проєкті, які дозволили реалізувати базовий функціонал програмного рішення. Тестування веб-додатку здійснювалося у напрямках тестування функціональності, відповідності, навантаження.

VIII. КОНФЛІКТИ ІНТЕРЕСІВ

Стосовно наведеної у наданих матеріалах інформації конфлікти інтересів, пов'язані з публікацією, відсутні.

ЛІТЕРАТУРА

- [1] P. Fidalgo, J. Thormann, and J. Lencastre, "Students' perceptions on distance education: A multinational study", *International Journal of Educational Technology in Higher Education*, vol. 17, no. 18 pp. 11–15, 2020. doi:10.1186/s41239-020-00194-2
- [2] A. Stevanovich, R. Bozshich, and S. Radovich, "Higher education students' experiences and opinion about distance learning during the Covid-19 pandemic", *Journal of Computer Assisted Learning*, vol. 6, no. 3, pp. 1683-1692, 2021. doi:10.1111/jcal.12613
- [3] Y. Chandra, "Online education during COVID-19: Perception of academic stress and emotional intelligence coping strategies among college students", *Asian Education and Development Studies*, vol. 10, no. 2, pp. 229–238, 2021. doi:10.1108/AEDS-05-2020-0097
- [4] I. Goksu, N. Ergun, Z. Ozkan, and H. Sakız, "Distance education amid a pandemic: Which psycho-demographic variables affect students in higher education?", *Journal of Computer Assisted Learning*, vol. 37, pp. 1539-1552, 2021. doi:10.1111/jcal.12544
- [5] H. Tseng, and E. J. Walsh, "Blended vs. traditional course delivery: Comparing Students' motivation, learning outcomes, and preferences", *Quarterly Review of Distance Education*, vol. 17, no. 1, pp. 11–14, 2016.
- [6] S. Vanslambrouck, C. Zhu, K. Lombaerts, B. Philipsen, and J. Tondeur, "Students' motivation and subjective task value of participating in online and blended learning environments", *Internet and Higher Education*, vol. 36, pp. 33–40, 2018. doi:10.1016/j.iheduc.2017.09.002
- [7] M. Wotto, "The future high education distance learning in Canada, the United States, and France: Insights from before COVID-19 secondary data analysis", *Journal of Educational Technology*, vol. 46, no. 2, pp. 262–281, 2020. doi:10.1177/0047239520940624
- [8] T. R. Padmanabhan, *Programming with Python*, Singapore: Springer, 2016, 344 p.

Пілотна система енергетичного моніторингу для навчальної будівлі університету

Андрій Яворський¹, Максим Карпаш², Віталій Цих¹, Любомир
Жовтуля²

¹Кафедра енергетичного менеджменту і технічної діагностики
¹Івано-Франківський національний технічний університет нафти і газу
²Університет Короля Данила
Івано-Франківськ, Україна

Анотація—Будівлі є одними з найбільших споживачів енергії. Відповідно, найбільше енергії споживається у великих громадських будівлях, до яких відносяться і приміщення університетів. Управління потоками енергії в таких будівлях і забезпечення їх енергоефективності починається з контролю та аналізу витрати енергії. Представлено реалізацію пілотної системи енергетичного моніторингу університетської будівлі на прикладі навчального корпусу Івано-Франківського національного технічного університету нафти і газу. Дана система моніторингу побудована на базі пристроїв smart-МАС, що дозволяють вимірювати споживання електричної та теплової енергії, холодної і та гарячої води, параметрів мікроклімату в приміщеннях будівлі та метеопараметрів, які впливають на споживання енергії. Система енергетичного моніторингу на базі пристроїв smart-МАС легко розгортається і налаштовується, приспосовування пристроїв для передачі даних здійснюється через мережу Wi-Fi, яка вже наявна в більшості будівель. Всі зібрані дані в системі зберігаються в хмарному сховищі з хвилинною деталізацією і доступні для подальшого аналізу та прийняття рішень у зручному WEB-додатку.

Ключові слова—Система Енергетичного Моніторингу Будівлі, Енергоспоживання, Облік Енергоносіїв, Показники Мікроклімату.

I. ВСТУП

Безперервний моніторинг споживання тепла, електричної енергії та води в громадських будівлях є основою для ефективного управління енергозабезпеченням та енергоспоживанням. У реальних умовах експлуатації будівель визначення їх поточного енергоспоживання за різними видами енергоресурсів проводиться за рахунок використання пристроїв обліку – лічильників тепла, електроенергії та води. Існуюча практика оперує лише місячними показами енергоспоживання будівель, що надзвичайно утруднює процес оперативного аналізу та, відповідно, ускладнює процес прийняття рішень щодо зменшення енергоспоживання при дотриманні нормативних умов мікроклімату у приміщеннях будівель. Таким чином, актуальним є розроблення автоматизованої системи моніторингу та управління енергозабезпеченням, що має реалізовувати завдання визначення годинних та добових показників енергоспоживання різними муніципальними будівлями за різними видами енергоресурсів з фіксацією параметрів внутрішнього мікроклімату, а також здійснювати управління функціонуванням системи енергозабезпечення на основі оперативного прийнятих рішень.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

На будівлі припадає біля 40% світового споживання енергії, що спричиняє біля третини усіх викидів парникових газів [1]. Витрати енергії для будівель в Україні є у 2–3 рази вищими, ніж подібні витрати в країнах ЄС. Практично 90% будівель України не відповідають сучасним вимогам енергоефективності, що призводить до перевитрат енергії на


опалення та недотримання вимог по мікроклімату приміщень [2]. Ці будівлі включають в себе значну кількість громадських будівель – від дитячих садків та шкіл, і до університетських, адміністративних будівель та закладів охорони здоров'я і культури. Відповідно, на утримання цих будівель в Україні витрачається значна частина коштів з місцевого та державного бюджетів [3].

На даний час на ринку присутні досить багато рішень для енергетичного моніторингу будівель [4-8]. Проте саме для громадських будівель, враховуючі значні площі і велику кількість точок вимірювання енергетичних параметрів і параметрів мікроклімату, важливою є простота і вартість побудови системи, інтеграція з існуючими засобами обліку, гнучкість і доступність у налаштуванні візуалізації та зберігання отриманих даних, що не потребують спеціалізованих навичок.

Науковцями Івано-Франківського національного технічного університету нафти і газу (ІФНТУНГ) було проведено ґрунтовне дослідження щодо пошуку найбільш оптимального рішення для побудови системи енергетичного моніторингу університетської будівлі, яка має відповідати наведеним вище вимогам. Для побудови системи моніторингу обрано «smart»-пристрої енергомонітори української компанії smart-MAIC [9].

Модельний ряд пристроїв smart-MAIC представлений енергомоніторами для постійного вимірювання параметрів електричної мережі та споживання електроенергії, універсальними імпульсними лічильниками для вимірювання споживання води, газу, тепла і за допомогою відповідних датчиків вимірювання температури, вологості, тиску, концентрації вуглекислого газу в повітрі, швидкості та напрямку вітру та інших параметрів (рисунок 1).

ПРОДУКТИ smart-MAIC





<div style="display: flex; align-items: center;">  <div style="margin-left: 20px;"> <p>Енергомонітор D101 (1 лінія) D103 (3 ліній)</p> </div> </div> <p style="margin-top: 20px;">Моніторинг та аналіз електроенергії</p> <p>Основні характеристики</p> <ul style="list-style-type: none"> напруга - V струм - A активна потужність - W реверсивна активна потужність - rW активна енергія - Wh зворотна активна енергія - rWh коефіцієнт потужності - cosФ вартість від 2390 грн. 	<div style="display: flex; align-items: center;"> <div style="margin-right: 20px;"> <p>Універсал D105</p> </div>  </div> <p style="margin-top: 20px;">Облік електроенергії, води, газу, тепла ...</p> <p>Температура, вологість, тиск, CO2, TDS, pH ...</p> <p>Основні характеристики</p> <ul style="list-style-type: none"> 2 імпульсних входи 1 аналоговий вхід (0-15В) 5 датчиків температури 1 універсальний датчик RS485 на вхід та вихід живлення 220В або 5В вартість від 990 грн.
--	--

Рисунок 1 – Модельний ряд пристроїв smart-MAIC

Пристрої smart-MAIC легко встановлюються і підключаються. Всі пристрої виконані в універсальному корпусі з можливістю монтажу на DIN рейку. Після включення пристрій доступний як точка доступу Wi-Fi, початкове налаштування можливе з будь-якого мобільного пристрою і займає лічені хвилини. Спеціальні налаштування для роботи не потрібні, пристрій самостійно приєднується до хмарного сервера даних і забезпечує двонаправлений обмін інформацією. Після встановлення і початкового налаштування пристрій smart-MAIC почне перетворення виміряних параметрів в інформацію, яка буде відправлена за допомогою бездротової технології Wi-Fi на хмарний сервер даних. Для аналізу і візуалізації даних використовується універсальний WEB-додаток smart-MAIC Dashboard, який є в звичайному інтернет браузері, а також як додаток для платформ

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

Windows, Android і iOS. Всі зібрані дані зберігаються на хмарному сервері з хвилинною деталізацією. При роботі в режимі реального часу показання з лічильників оновлюються з інтервалом 5 секунд. Smart-MAIC Dashboard дає можливість контролювати поточні показання і візуалізувати історичні дані отримані з пристроїв smart-MAIC, наприклад, для оптимізації споживання при двотарифному або багатотарифному обліку. Користувачеві доступні гнучке налаштування віджетів індикаторів і графіків, необмежена кількість сторінок і підключених пристроїв до одного облікового запису. Так само поточні дані доступні на WEB сторінці самого пристрою.

III. РЕЗУЛЬТАТИ

Пілотна системи енергетичного моніторингу для університетської будівлі з використанням пристроїв smart-MAIC інсталювана на базі кафедри «Енергетичного менеджменту і технічної діагностики» (ЕМіТД) ІФНТУНГ. Кафедра ЕМіТД займає третину будівлі навчального корпусу №9 університету і розміщується на двох поверхах будівлі (рисунок 2). Відповідно, на основі приміщень кафедри можна повністю змодельовати систему енергетичного моніторингу, яку потім можна буде за аналогією поширювати на інші будівлі університету.



Рисунок 2 – Навчальний корпус №9 ІФНТУНГ (а) та аудиторії кафедри «Енергетичного менеджменту і технічної діагностики» (б)

Структура пілотної системи енергетичного моніторингу для навчальної будівлі університету наведена на рисунку 3. Система енергетичного моніторингу дозволяє отримати реальні дані по споживанню електричної та теплової енергії, холодної та гарячої води для університетської будівлі. Окрім того в кожному навчальному приміщенні передбачено моніторинг основних показників мікроклімату – температури і вологості повітря, концентрації вуглекислого газу. Окремою складовою системи моніторингу є метеорологічний модуль, що дозволить відслідковувати температуру і вологість зовнішнього повітря, напрямок і швидкість вітру, атмосферний тиск, інтенсивність сонячного випромінювання (інсоляцію), значення і тренд зміни атмосферного тиску.

Як було зазначено, продуктова лінійка smart-MAIC представлена рішеннями для моніторингу будь-яких подій, станів і процесів – однофазний електролічильник та трифазний електролічильник з кільцевими або роз'ємними трансформаторами струму, лічильники імпульсів з давачами температурних режимів і аналоговим входом. Пристрої мають зручне виконання у вигляді корпусу на DIN-рейку відповідно до BS EN 60715:2017, що дозволяє монтувати їх у будь-якому новому або вже змонтованому електрошриту. На рисунку 4 наведено приклад інсталяції сегменту системи енергетичного моніторингу для навчальної аудиторії університету.

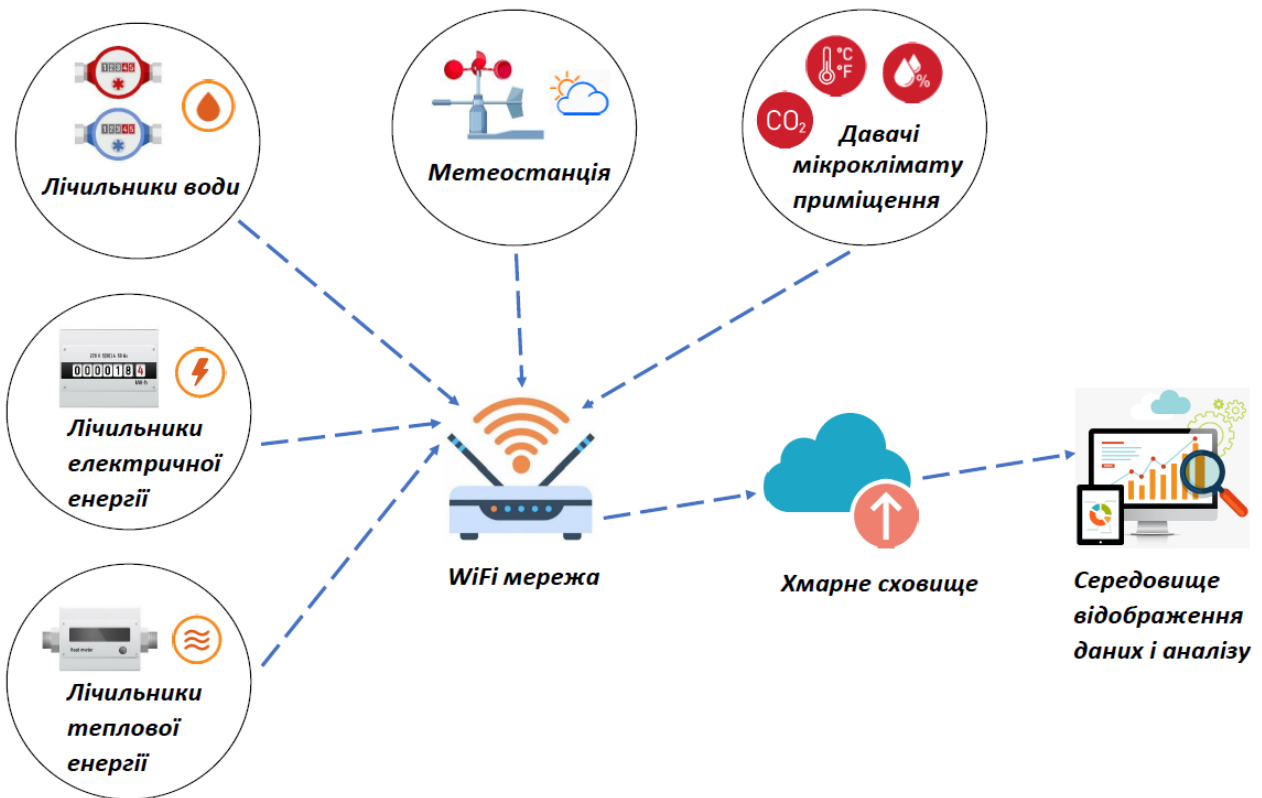
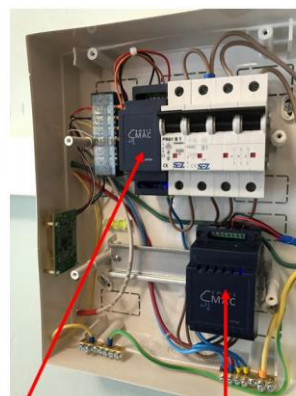


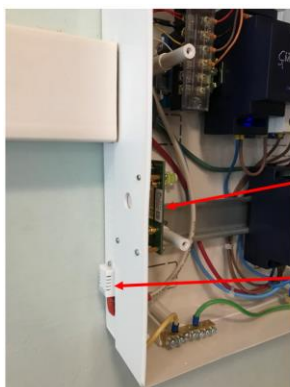
Рисунок 3 – Загальна структура пілотної системи енергетичного моніторингу для навчальної будівлі університету



Універсальний монітор D105



Енергомонітор D103



CO₂ давач (MH-Z14A)

Давач температури і вологості (DHT22)



Трансформатори струму (100А) для енергомонітора D103

Рисунок 4 – Приклад інсталяції сегменту системи енергетичного моніторингу для навчальної аудиторії університету

В навчальній аудиторії за допомогою системи енергетичного моніторингу відслідковуються параметри мікроклімату в приміщенні (температура, вологість і рівень вуглекислого газу) для цього використовується універсальний монітор smart-MAIC D105 до якого приєднується датчик температури-вологості (DHT22) і вуглекислого газу (MH-Z41A). Для визначення рівня споживання електричної енергії в аудиторії (навчальна аудиторія обладнана електричним опаленням) використовується енергомонітор smart-MAIC D103, що забезпечує приєднання до трифазної лінії електропостачання по напрузі і по струму (за допомогою кільцевих трансформаторів струму на номінал 100А). Дані, які збираються з пристроїв, зберігаються в хмарному сховищі і доступні для перегляду та аналізу в режимі реального часу. Агрегування інформації, щодо конкретної аудиторії чи інших об'єктів моніторингу, її візуалізація та подальший аналіз здійснюється в хмарному WEB-додатку smart-MAIC Dashboard. Користувач сам налаштовує вид віджетів індикаторів, графіків та таблиць для інформаційної панелі. На рисунку 5 наведений приклад інформаційної панелі сегменту системи енергетичного моніторингу для навчальної аудиторії університету.



Рисунок 5 – Приклад інформаційної панелі сегменту системи енергетичного моніторингу для навчальної аудиторії університету

На інформаційну панель (рисунок 5) виводяться поточні значення температури, вологості повітря та концентрації вуглекислого газу в приміщенні, напруги, струму, споживаної потужності по кожній фазі лінії електропостачання. Важливим для аналізу енергоспоживання, температурної інерції будівлі, ефективності роботи системи опалення кондиціонування та вентиляції є відслідковування зміни контрольованих параметрів в часі. WEB-додаток smart-MAIC Dashboard дає можливість побудови різноманітних графічних залежностей з різною часовою деталізацією (хвилина, година, день, тиждень, місяць, рік) і в з різним типом графічного відображення трендів (лінія, площа, стовпчикова діаграма). Додатково користувач може сконфігурувати виведення необхідних даних у вигляді таблиці, зокрема для деталізації споживання та витрат, задавши відповідно вартість для енергоносія чи гарячої/холодної води, витрата яких відслідковується системою моніторингу. Для подальшого аналізу, до прикладу, в середовищі Excel, дані агреговані в таблиці експортуються в форматі csv.

Для оцінки споживання електричної енергії усією будівлею чи її частиною теж використовується енергомонітор smart-MAIC D103 з відповідними трансформаторами струму. В лінійці пристроїв smart-MAIC використовуються трансформатори струму від 100А до 1000А, що дозволяє контролювати навантаження в будівлі до 1МВт. Приклад інсталяції сегменту системи енергетичного моніторингу для визначення споживання електричної енергії частиною будівлі показано на рисунку 6. В даному випадку моніториться споживання електричної енергії для цілої кафедри, що займає частину будівлі. Відповідно використовується енергомонітор smart-MAIC D103 з накладними кільцевими трансформаторами струму на 300А. Для візуалізації та аналізу даних по споживанню електричною енергією кафедрою формується окрема інформаційна сторінка в хмарному WEB-додатку smart-MAIC Dashboard аналогічна до наведеної на рисунку 5.



Рисунок 6 – Приклад інсталяції сегменту системи енергетичного моніторингу для визначення споживання електричної енергії частиною будівлі

З метою комплексного вирішення задач енергетичного моніторингу прилади сторонніх виробників теж інтегруються з пристроями smart-MAIC – лічильники води, тепла, палива, давачі до газових лічильників та інші витратомірів з імпульсними виходами. Важливим є те, що такий підхід дозволяє задіяти в систему енергетичного моніторингу прилади обліку, що вже наявні в будівлі. В цьому випадку використовується універсальний монітор smart-MAIC D105, який обладнаний двома імпульсними входами. На рисунку 7 показано приклад реалізації моніторингу споживання гарячої води в будівлі і витрати електричної енергії на її приготування. В цьому випадку використовується лічильник гарячої води з імпульсним

видом, який змонтований на виході електричного бойлера і класичний електронний лічильник з механічним обліковим пристроєм (через який відбувається електроживлення бойлера), що має телеметричний імпульсний вихід. Відповідно ці лічильники приєднуються до універсального монітора smart-MAIC D105 для передачі даних в систему енергетичного моніторингу будівлі.

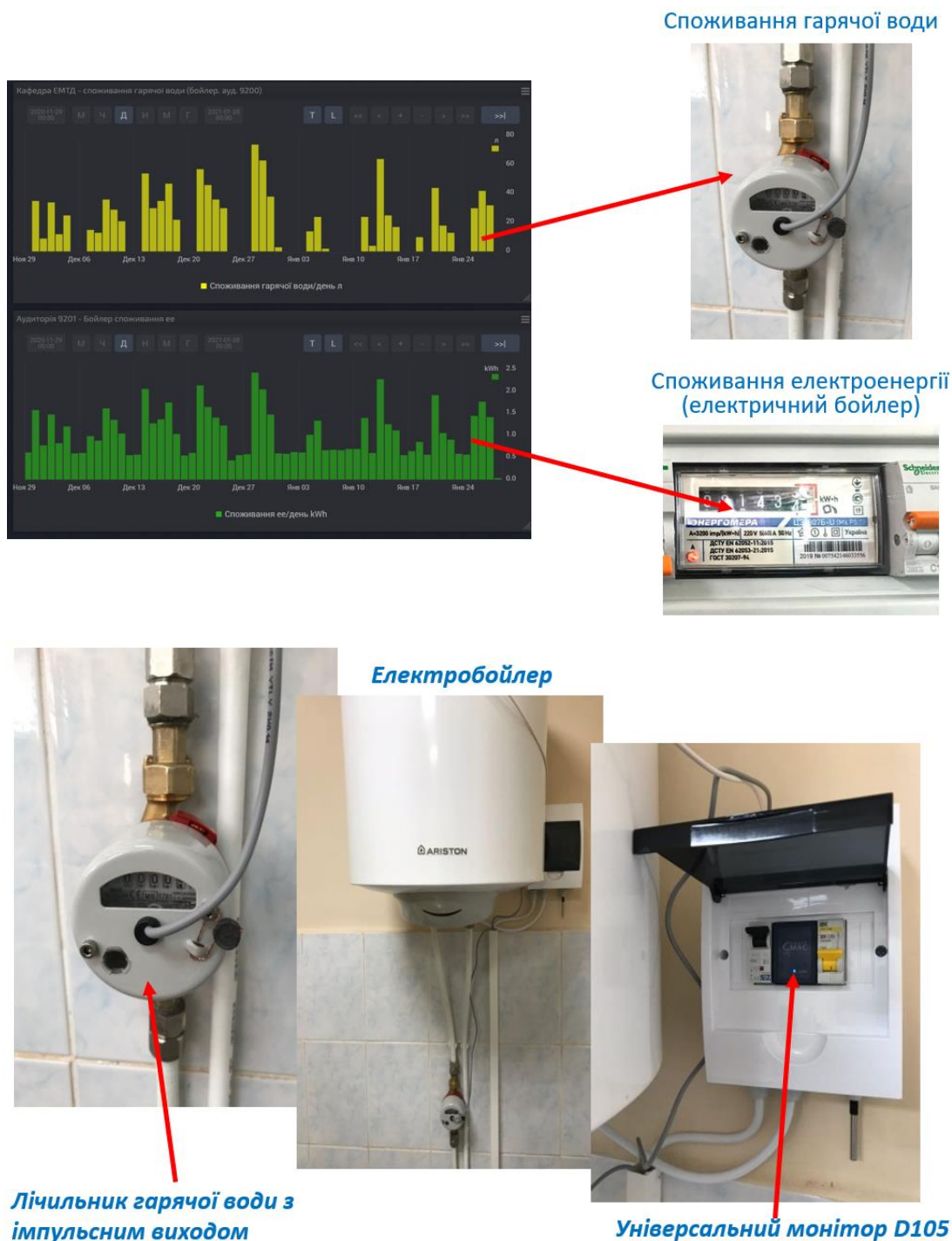


Рисунок 7 – Приклад реалізації моніторингу споживання гарячої води в будівлі і витрати електричної енергії на її приготування

Обов'язково в системі енергетичного моніторингу будівлі має бути метеорологічний модуль для отримання локальних метеоданих. Адже для оцінки енергетичної ефективності будівлі необхідно володіти інформацією щодо споживання теплової енергії або витрати природного газу чи електроенергії, що використовуються на потреби опалення будівлі, в

залежності із зовнішньою температурою повітря. На рисунку 8 наведено реалізацію частини метеорологічного модуля системи енергетичного моніторингу будівлі, що призначений для реєстрації температури та вологості зовнішнього повітря і значення атмосферного тиску. В якості первинного перетворювача використовується датчик температури, вологості і атмосферного тиску BMP280, який приєднаний до універсального монітора smart-MAIC D105. Для забезпечення отримання коректних метеоданих датчик змонтований в радіаційному екрані, для захисту від сонячного випромінювання, і розміщується на північній стороні будівлі в затінку. Також система енергетичного моніторингу дозволяє інтегрувати в себе і інші пристрої, наприклад ті, що відносяться до систем безпеки. На рисунку 8 наведено як в систему енергетичного моніторингу будівлі долучено дозиметр радіометр для контролю за станом радіаційного фону.



Рисунок 8 – Реалізація частини метеорологічного модуля системи енергетичного моніторингу будівлі, що призначений для реєстрації температури та вологості зовнішнього повітря і значення атмосферного тиску

IV. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В подальшому доцільно впровадити аналогічні системи моніторингу енергетичного споживання як в інших приміщеннях даного корпусу, так і в інших будівлях університету. Це дасть можливість чітко прослідкувати та по можливості регулювати навантаження в системі енергоспоживання.

V. ВИСНОВКИ

Отримані дані по фактичному енергоспоживанню, після розгортання пілотної системи моніторингу, протягом календарного року дозволять зокрема оцінити питомі теплові характеристики будівлі, визначити фактичну енергопотребу для забезпечення нормативних показників мікроклімату в приміщеннях будівлі, оцінити раціональність і тренди споживання енергоносіїв. Також такі дані допоможуть оцінити реальний ефект після впровадження енергоефективних заходів для зменшення споживання енергоресурсів в будівлях університету.

ЛІТЕРАТУРА

- [1] P. Nejat, F. Jomehzadeh, M. M. Taheri, M. Gohari and M. Z. A. Majid, “A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries),” *Renewable and sustainable energy reviews*, vol.43, pp. 843-862, 2015.
- [2] O. M. Shevchenko, M. M. Shovkalyuk, “KPI energy efficient campus: instruments and research methods”, *Bulletin of Kyiv National University of Technology and Design. Series: Technical Sciences*, vol. 4, pp. 97-105, 2019.
- [3] Y. Parfenenko, V. Shendryk, V. Nenja and S. Vashchenko, “Information system for monitoring and forecast of building heat consumption” in *Int. Conf. Information and Software Technologies*, Oct, 2014, pp. 1-11.
- [4] L. Zhao, J. L. Zhang and R. B. Liang, “Development of an energy monitoring system for large public buildings”, *Energy and Buildings*, vol. 66, pp. 41-48, 2013.
- [5] B. Baranyai and I. Kistelegdi, “Energy management monitoring and control of public buildings”, *Pollack Periodica*, vol. 9(2), pp. 77-88, 2014.
- [6] D. Ibaseta, A. García, M. Álvarez, B. Garzón, F. Díez, P. Coca and J. Molleda, “Monitoring and control of energy consumption in buildings using WoT: A novel approach for smart retrofit”, *Sustainable Cities and Society*, vol. 65, pp. 22-37, 2021.
- [7] B. Mataloto, D. Calé, K. Carimo, J. C. Ferreira and R. Resende, “3D IoT System for Environmental and Energy Consumption Monitoring System”, *Sustainability*, vol. 13(3), pp. 14-95, 2021.
- [8] V. K. Nguyen, W. E. Zhang and A. Mahmood, “Semi-supervised Intrusive Appliance Load Monitoring in Smart Energy Monitoring System”, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 17(3), pp. 1-20, 2021.
- [9] “Smart-MAIC. Smart meters for any kind of consumptions”. Smart-maic.com. <https://smart-maic.com/en/> (accessed Oct. 21, 2021).

Impact of Covid-19 on University Education Digitalization: A Critical Interpretive Synthesis of Tensions and Barriers for Automation Field

Iryna Tarasevich, Alena Tarasevich, Mykhailo Kuzmenko

*Production Process Automation Department
Institute of Industrial and Business Technologies
Ukrainian State University of Science and Technologies
Dnipro, Ukraine*

*Centre for Education Research and Innovation (CERI)
The University of Western Ontario
London, Ontario, Canada*

Abstract—What did Covid-19 bring into university digitalization? A succinct and all-embracing answer would be “a lot of different things on many different levels”. Automation and digitalization, hereinafter collectively referred to as technology, have been around for a while, although never used to the full potential of their in-built, inscribed features. The pandemic was a huge stimulator to explore new avenues in university digitalization and this critical interpretive synthesis is aimed at developing theory of major tensions and barriers that main users of technology or agents (such as students, faculty members and university administration) have been continuously challenged with. The tensions and barriers described below are presented from three distinct perspectives – from the perspectives of students, faculty members, and university administrators. Whether students thrive or survive, whether the faculty rethink and upskill their teaching, whether the university as an institution can potentially meet multiple, often competing needs of both, and whether all the parties involved are able to re-conceptualize their roles in the process of university automation and digitalization is yet to be thoroughly studied by future research.

Keywords—*University Education, Digitalization, Automation, Pandemic.*

I. INTRODUCTION: DIGITALIZATION OF UNIVERSITY EDUCATION IN A STATE OF EMERGENCY

Although in many disruptive, yet powerful ways, nothing stimulates the humankind for development better than a crisis – be it a pandemic, a revolution, a war, or any other major natural disaster or a socioeconomic change in the human history. This has been the case with the wide-reaching impact of Covid-19 on university digitalization that brings in its associated tensions and challenges – both pre-existing and Covid-specific. Like many areas of educational practice, the literature on education digitalization and digital universities is large, complex, and diverse and includes such various types of evidence, as quantitative and qualitative studies, theoretical and methodological work, as well as epidemiological studies. With the pandemic having been a ‘hot shot’ for several years now, many relevant articles are being published on a daily basis. That is why it would be beneficial to conduct a review of the area that could produce a “mid-range” theoretical account of university digitalization under a crisis that can be applied to different educational settings [1].

II. RELATED WORK

Critical interpretive synthesis (CIS) is a relatively new type of review methodology that synthesizes different types of evidence to critically develop new theory of a phenomenon. Originated in health equity research in 2006 [1], CIS was initially uptaken in health professions education and since 2013 other social sciences studies have begun using it in their reporting practices [2]. Sensitized to both conventional systematic review methodology and qualitative methods of interpretative synthesis, CIS does not prioritize quantitative evidence over qualitative inquiry and assumes that both have something to add to theory building in a topic area. While a conventional systematic review is beneficial for appraising and summarizing evidence from primary studies [3] with the aim of testing theories and finding out “what works” (aggregative synthesis), it is more limited when the aim of a review is to develop theories (interpretive synthesis) [1]. Contrary to the conventional systematic review methodology with its clear account of pre-defined and reproduced searching strategy [3], CIS relies on a more organic review process. It fits better the emergent nature of a review question and combines several search strategies. CIS provides not a definitive answer, but rather important insights into human experiences of a phenomenon, acceptability of policy, and a deeper understanding about why something might or might not work. It can also open up a floor for future inquiries into aspects of theory that have long been taken for granted and need to be problematized for their resolution. When carefully executed, CIS is also useful for synthesizing literature in preparation of a research grant or a practice project.

III. METHODOLOGY

The aim of this review is to construct a critical analysis of a complex and diverse body of literature on university digitalization during the COVID-19 pandemic. Although the two approaches described above (aggregative and interpretive syntheses) serve their own goals of theory developing or testing, for the purpose of exploring tensions and barriers to university digitalization in the pandemic context, interpretive synthesis seems to be more relevant. Given the relative novelty and diversity of university digitalization under Covid, it is neither desirable, nor feasible to formulate a review question – accurately and a priori. First, a tightly defined review question would make it difficult for new insights to emerge from the literature review. Second, the aim is to broadly explore tensions and barriers to university digitalization, with a focus on theory building on the topic, rather than on testing a specific hypothesis. For this purpose and contested at the outset of the study, the review question has to be tentative and fuzzy, emergent and exploratory [1].

A compass rather than an anchor [1], our research question is explored during the review process around three major tenets: 1) tensions and barriers to university digitalization from multiple perspectives of its main users or agents (students, faculty members and university administration); 2) how pandemic reality (the enacted) impacts the inscribed features of technology (the espoused); 3) and what lessons can be learnt from this experience in order to meaningfully advance university digitalization in the future. Guided by an iterative approach to review and informed by emerging insights from the literature, the review question has been answered along the review process. The multidisciplinary of the reviewers is also a benefit, allowing a range of perspectives to be incorporated into different review stages.

In the literature search, we first retrieved relevant papers dated 2019-2022 from two electronic databases, the Web of Science and Scopus. The database search was then complimented with reference chaining, expert contacts and using the team expertise. As a result, some of the retrieved papers were not explicitly related to university digitalization, per se, yet important to the aim of the review [1]. Given the already diverse, large body of literature on university digitalization during the pandemic, the aim of the searching phase was shifted from including all the relevant literature (conventional systematic review) with a broad, exhaustive search strategy (university* education* COVID* key words in the topic field that searches titles, abstracts, and author keywords) to developing a sampling frame of some relevant papers that were identified with a narrower search

strategy (university* digitalization* COVID* key words). Therefore, for practical reasons, we moved away from 14,004 potentially relevant papers retrieved from Scopus and the Web of Science to 380 relevant records.

At the initial stages of the review, purposeful sampling was used to select papers that are relevant to the review purpose, i. e., information-rich papers that directly address barriers and challenges of university digitalization and their impact in the titles, abstracts, or authors' keywords. Later in the review process, theoretical sampling, i. e., sampling for the purpose of theory development, guided data collection and analysis. In other words, as the theory of COVID-19 impact on university digitalization was emerging from the literature, we decided which data to collect next and where to find them [4].

IV. FINDINGS

As a result of this critical interpretive synthesis of literature, some key challenges and barriers to university digitalization were identified from the perspectives of main agents: students, professors, and university administrators. However, the most challenging part of university digitalization that all three groups of agents share is to overcome people's resistance to change (an entirely human factor). As a primary and most powerful barrier, a human factor is always at play and relates to knowledge, skills and culture [5], shaping any automation process. Furthermore, emergency remote teaching reveals a number of human factors that have been largely ignored by university education: lack of motivation, socio-emotional distance, socio-economic gaps and cultural isolation [6]. In an attempt to balance technology, pedagogy and the 'new normal', agents involved in university education are forced to be proactive and productive in their work while developing new skills fast in a stressful environment.

From students' perspective, the major challenge is technology [7] that in itself can further amplify the digital divide for vulnerable student populations [8]. Some also report it difficult to stay focused [7] while others suffer from lack of self-management and time to follow different subjects. Moreover, in a purely online context of self-isolation, the feeling of competitiveness as a powerful stimulator of learning can be missing. Whether students thrive or survive or are situated in-between is still to be explored by future research.

Professors also report isolation as a barrier to designing courses. The rapid transition to online teaching left some educators grappling with the necessity to redesign their courses to modalities they might have never taught in. This was particularly problematic at the beginning of the pandemic when they had no pedagogies for online teaching and evaluation [9]. With little training in some cases and in no time, the faculty were pressured to adapt quickly to new online methodology and acquire new competencies in technology use. Given the diverse levels of preparedness to use technology by faculty members (a digital divide), it is also reasonable to add a generational divide [8] to a list of powerful barriers and challenges to automation. What remains unresolved, however, and is particularly problematic for training students in the automation subspecialty, specifically, is providing feedback and ongoing support to students. Mainly dependent on laboratory equipment and IT-infrastructure, automation students, while on remote learning, cannot receive timely feedback from their instructors. Oftentimes, the process of saving, converting, and sending huge chunks of electronic control packages to their instructors for review is so time-consuming that feedback no longer serves its purpose. This is the case with any subspecialty that heavily relies in its training on physical, technological resources that are available only on campus and cannot be substituted otherwise.

At the institutional level, universities deal with similar tensions and barriers while also being challenged by financial constraints and limitations of their in-built IT infrastructures [10]. More than anything, such human factors, as lack of IT knowledge and skills to apply automation in specific contexts, pose a major challenge for administration. Similar to faculty members, administrators are expected to rapidly acquire skills in using new technologies while still grappling with a lack of culture for new modes of communication [5].

V. DISCUSSION

The pandemic has demonstrated the necessity of rethinking roles in university education digitalization. It has shown great potential of automation and its biggest challenge of overcoming material and social dimensions of traditional educational practices. Socially, organizationally and materially, there are multiple digital learning spaces that can overlap with hybrid and physical spaces. For example, within a matter of minutes and almost effortlessly, a student can move from an in-person to a virtual lecture, from Wikipedia articles to an industry webinar, from an international discussion group to social media. All these ‘platforms’ are valid, legitimate ways of teaching and learning and that is why some farseeing researchers have long argued that future university digitalization will entail outside agents to be included into the process [11, 12]. In the years to come, this new experience can help bridge the gap between in-person and online learning. University digitalization, as experienced and lived through under Covid, is dual in nature because we clearly deal with digitalization of education and digitalization of subjects, i. e., all human agents involved in the process (students, faculty members, administrators, technical support staff, etc.).

VI. FUTURE RESEARCH

The final account of the impact COVID-19 had had on university education, was shaped later in the review process. Moreover, much of the subsequent sampling was directed at developing and purposively challenging the theory of an impact as it began to emerge from the critical interpretive synthesis of the review findings. As mentioned above, such sampling methods and iterative nature of theory development are neither reproduced nor audited. Testing whether interpretations of COVID effects on university education change with new evidence, will be an important avenue to pursue in future research. It will also be interesting to assess further uptake of critical interpretive synthesis methodology and its potential in re-conceptualizing complex phenomena in quantitative and qualitative empirical work.

VII. CONCLUSION

The critical interpretive synthesis of the literature of the pandemic impact on university digitalization reveals a number of challenges and barriers along the process from the perspectives of three main groups of agents: students, professors, and administrators. This summary work is timely and important as emerged (and emerging) digital learning spaces are here to stay and will likely coexist with traditional delivery modes of university education. Although the importance of physical spaces and technology/laboratory resources of universities is hard to underestimate, remote teaching and learning will continue to thrive. That is why it is so important to find a meaningful balance between the two and develop new participatory culture that is shared by the key agents and supports automation changes.

VIII. DISCLOSURES

The authors declare that they have no conflicts of interest related to this paper.

REFERENCES

- [1] M. Dixon-Woods, S. Bonas, A. Booth, D. R. Jones, T. Miller, A. J. Sutton, R. L. Shaw, J. A. Smith, and B. Young, “How can systematic reviews incorporate qualitative research? A critical perspective,” *Qual. Res.*, vol. 6, no. 1, pp. 27–44, 2006, doi: 10.1177/1468794106058867.
- [2] J. Depraetere, C. Vandeviver, I. Keygnaert, and T. V. Beken, “The critical interpretive synthesis: An assessment of reporting practices,” *Int. J. Soc. Res. Methodol.*, vol. 24, no. 6, pp. 669–689, 2021, doi: 10.1080/13645579.2020.1799637.
- [3] M. Egger, G. Davey Smith, and A. Phillips, “Meta-analysis: Principles and procedures,” *BMJ*, vol. 315, 1533–1537, 1997.

- [4] I. T. Coyne, "Sampling in qualitative research. Purposeful and theoretical sampling; Merging or clear boundaries?" *J. Adv. Nurs.*, vol. 26, no. 3, pp. 623–630, 1997.
- [5] E. Stefanova, "Digital transformation catalyzed by Covid-19 pandemic: Challenges and solutions," *Pedagogica*, vol. 94, no. 3s, pp. 13–24, 2022, doi: 10.53656/ped2022-3s.01.
- [6] C. Rapanta, L. Botturi, P. Goodyear, L. Guàrdia, and M. Koole, "Balancing technology, pedagogy and the new normal: Post-pandemic challenges for higher education," *Postdigit. Sci. and Educ.*, vol. 3, no. 3, pp. 715–742, 2021, doi: 0.1007/s42438-021-00249-1.
- [7] L. Mishra, T. Gupta, and A. Shree, "Online teaching-learning in higher education during lockdown period of COVID-19 pandemic," *IJEDRO*, vol.1, no. 100012, 2020, doi: 10.1016/j.ijedro.2020.100012.
- [8] V. Govindarajan and A. Srivastava, "What the shift to virtual learning could mean for the future of higher education," *Harv. Bus. Rev.* Accessed: Sept. 29, 2022. [Online]. Available: <https://hbr.org/2020/03/what-the-shift-to-virtuallearning-could-mean-for-the-future-of-higher-ed>
- [9] C. Carolan, C. L. Davies, P. Crookes, S. McGhee, and M. Rox-Burgh, "Covid-19: Disruptive impacts and transformative opportunities in undergraduate nurse education," *NEP*, vol. 46, no. 102807, 2020, doi: 10.1016/j.nepr.2020.102807.
- [10] S. Krishnamurthy, "The future of business education: A commentary in the shadow of the Covid-19 pandemic," *J. Bus. Res.*, vol. 117, pp. 1–5, 2020, doi: 10.1016/j.jbusres.2020.05.034.
- [11] R. Hazemi, S. Hailes, and S. Wilbur, Eds. *The Digital University: Reinventing the Academy*. Springer Science & Business Media, 2012.
- [12] N. C. Jackson, "Managing for competency with innovation change in higher education: Examining the pitfalls and pivots of digital transformation," *Bus. Horiz.*, vol. 62, no. 6, pp. 761–772, 2019, doi: 10.1016/j.bushor.2019.08.002.

Selection of data preprocessing technique for imbalanced dataset

Olga Solovei

*Department of Information Technology and Applied mathematics
Kyiv National University of Construction and Architecture,
Ukraine*

Abstract—In many domains datasets are imbalance, most often the problem is visible in datasets with a medicine diagnoses' prediction; credit's approval; surveys. The results of machine learning algorithms for imbalanced dataset are biased to majority class, whereas the correct predictions of the smaller class somethings could be most important. The goal of the work is to empirically study which data preprocessing technique: feature selection or oversampling to be applied before to train a classification learning algorithm.

Keywords—*Oversampling, Feature Selection, Area under ROC curve, Random Forest classifier.*

I. INTRODUCTION

In study [1] to improve classification models' results for imbalance datasets are proposed the following data preprocessing strategies: 1) oversampling - the smaller class is oversampled to match the size of the major class; 2) down-sizing – the number of samples in the major class is reduced; 3) learning by recognition - to use unsupervised machine learning algorithms and ignore target class in dataset. From the mentioned three strategies the preference is given to oversampling technique as it makes balance distribution between classes which improves that classifier's performance. However, in study [2] can be found that the most widely used oversampling method - SMOTE is good for a highly dimensional imbalance dataset only when k-nearest neighbors classifier is used and most predictive features are selected.

Because the results of the mentioned studies look controversial more empirical tests are required to understand the optimal approach to preprocessing technique selection.

II. RELATED WORK

In work [3] was proposed to use only feature selection technique for a highly dimensional imbalance gene dataset without comparing its efficiency to the results of the oversampling or down-sizing, therefore the concluded results of the study are not obvious.

To eliminate that gap, in the current work, will be conducted experiments to evaluate the efficiency of data preprocessing techniques: 1) a feature selection; 2) an oversampling by SMOTE; 3) a combination of both.

III. METHODOLOGY

Our datasets are from UCI Machine Learning Repository: "adult", "coli" and "songs" are highly dimensional with imbalance ration between the minority to the majority class: 0.3, 0.06, 0.16 correspondingly. The features' count per each dataset is 81, 86 and 58 and samples' count per each dataset is 48842, 9822, 1409 correspondingly. In the experiments are included the steps: 1) create for each dataset two pairs of the train and test data using different data preprocessing techniques: pairset1 – the dataset is left imbalance but the most predictive features are selected by applying mutual information method [4]; pairset2 - the dataset is balanced by SMOTE method so that a minority class 100% matches the majority and the most predictive features are selected; pairset3 - the dataset is balanced by SMOTE method but all features are included. 2) train Random Forest

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

(RF) algorithm and evaluate the prediction quality using metrics: area under ROC (AUC), precision, recall and f1-score on the test subset; 4) compare performance's results based on the metrics' values and ROC curves' graphs.

IV. RESULTS

Random Forest classifier's performance metrics for each dataset are recorded in table 1. ROC convex hull graphs presented on fig 1.

Table 1. RF algorithm's performance results and training time

dataset	Pairset1				Pairset2				Pairset3			
	AUC, %	Precision, %	Recall, %	f1 score, %	AUC, %	Precision, %	Recall, %	f1 score, %	AUC, %	Precision, %	Recall, %	f1 score, %
adult	84	64	51	56	93	84	86	85	93	85	86	85
coli	65	52	51	51	99	96	95	96	99	97	95	96
songs	47	57	54	55	85	81	80	80	90	83	83	83

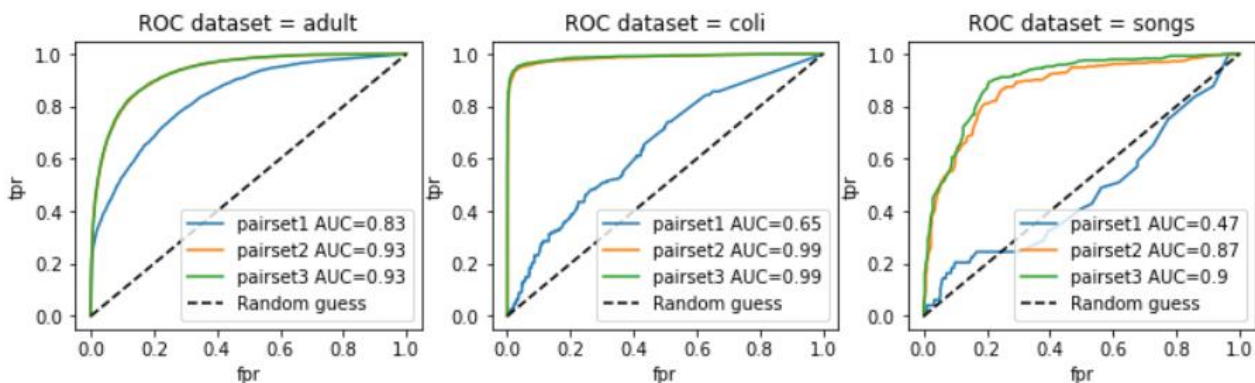


Figure 1. ROC curves of RF classifier trained on pairset1, partset2 and pairset3.

V. DISCUSSION

The results from table 1 can be interpreted as the following: for highly dimensional datasets an ability of RF classifier to predict the target class is better when a dataset is balanced (pairset2, pairset3) compared to the performance of RF on the imbalance dataset with the most predictive features (pairset1). However, for dataset = "coli" the performance of RF on the balance dataset with all features (pairset3) is better compared to the performance of RF on the balance dataset with the most predictive features are selected (pairset2).

The same results are visible on ROCs convex hull graphs from fig 1: the "more northwest" ROC curve is seen for pairset2 and pairset3 on datasets: "adult" and "coli" and the "more northwest" ROC curve is seen for pairset3 on dataset: "songs".

VI. FUTURE RESEARCH

It is worth to further investigate: 1) the patterns in the dataset when applying only feature selection technique can guarantee the high model's performance; 2) how model's performance is affected when down-sizing technique together with feature selection are applied on the data preprocessing stage.

VII. CONCLUSION

Random Forest classifier's performance is better when SMOTE method of the oversampling technique is applied compared to the performance's results received by applying dataset's dimension reduction by selecting the most predictive features using mutual information method.

REFERENCES

- [1] N. Japkowicz. "The class Imbalance Problem: Significance and Strategies," *AAAI conference of Artificial Intelligence*, 2000, pp.10-15.
- [2] R. Blagus, L. Lusa. "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinformatics*, 2013 – Springer, pp.2-16.
- [3] G.A.Sharifai, Z. Zainol. "Feature Selection for High-Dimensional and Imbalanced Biomedical Data Based on Robust Correlation Based Redundancy and Binary Grasshopper Optimization Algorithm," *MDPI Genes*, 2020, pp.-2-26.
- [4] J. Li, K. Cheng, S. Wang, F. Morstatter, R. Trevino, J. Tang & H. Liu. "Feature Selection: A Data Perspective," *ACM Computing Surveys*, vol. (50), no. 6, 2018, pp 1–45, DOI:<https://doi.org/10.1145/3136625>.

Дослідження можливості використання нейромережевих технологій для вирішення задачі класифікації дефектів металопрокату

Катерина Островська, Іван Стовпченко, Сергій Черський

*Кафедра інформаційних технологій і систем
Український державний університет науки та технологій
Дніпро, Україна*

Анотація - Робота присвячена розробці програмної системи для класифікації дефектів металопрокату на базі нейромережевих технологій. В роботі виконано дослідження предметної галузі, визначено можливість використання нейронних мереж для вирішення задачі класифікації дефектів металопрокату. Побудовано модель нейронної мережі, що підходить для вирішення задачі класифікації видимих дефектів металопрокату. Виконано генерацію навчальної, тестової та валідаційної вибірок на основі бази даних дефектів поверхні гарячого металопрокату. Спроектована та реалізована система класифікації дефектів.

Ключові слова - *Нейромережеві технології, Класифікація, Металопрокат, Валідаційна вибірка, База даних, Аугментація.*

I. ВСТУП

Контроль якості продукції є важливим етапом на металургійних заводах, адже гарантія якості означає відповідність механічних та інших властивостей певним значенням. Але коли метал перебуває у прокаті, стан його поверхні чи покриття створює окремий важливий чинник якості, який також необхідно контролювати [1].

Контроль поверхні зазвичай здійснюється візуально (людиною) на кінцевих етапах виробництва. Однак навіть тут швидкість прокату металу дуже висока і на проведення контролю є дуже мало часу. По закінченню процесу метал, що прокочується, згортається в рулон, після чого виявити дефект стає ще важче. Нерідко це призводить до відвантаження споживачеві неякісного металу, що надалі тягне за собою значні фінансові втрати для виробника.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Потреба металургійних підприємств у системах автоматизації виявлення та класифікації поверхневих дефектів є дуже високою. Але, незважаючи на цю високу потребу, розробників подібних систем на сьогоднішній день досить мало. Найбільш відомі: ISRA Parsytec, Codnex, Matra, EES, Siemens-VAI, Sipar та ін.) [2]. При цьому існуючі системи добре виконують виявлення дефектів, але досі мають складнощі з їхньою класифікацією. Складність обумовлена тим, що поверхні сталей, що прокочуються на різних агрегатах навіть одного і того ж заводу, часто різняться на вигляд, а це призводить до ускладнення алгоритмів виявлення та класифікації дефектів поверхні [3, 4, 5]. З іншого боку, будь-яку універсальну частину системи доводиться адаптувати під конкретну виробничу лінію [6]. Крім цього, складність обумовлена великою кількістю різновидів поверхневих дефектів [7].

Характеристики та класифікація дефектів варіюються для кожного конкретного виробництва та агрегату і не регулюються загальноприйнятим стандартом. Навіть невеликі зміни у виробничому процесі можуть спричинити появу нових видів дефектів [8].

Архітектура та докладні характеристики комерційних систем автоматичного виявлення дефектів не поширюються у відкритому доступі.

Наприклад, у роботах [9, 10] описаний приклад з використанням алгоритму виділення ознак, створеного без застосування машинного навчання у комбінації з нейронною мережею для класифікації.

Варто окремо відзначити, що висока точність класифікації дефектів створить можливість заздалегідь узгоджувати якість металопрокату зі споживачем – підприємство може піти на певні поступки, замість повної заміни та подальшого надвиробництва продукції.

Таким чином, розробка аналогічної системи з використанням штучних нейронних мереж (ШНМ) для досягнення високої точності класифікації є дуже актуальним завданням.

III. МЕТОДИ

Однією з головних складових процесу навчання ШНМ, а також якості її роботи надалі є набір вихідних даних – датасет.

Датасет дефектів гарячого металопрокату, що надається Північно-східним університетом США (NEU) спеціально призначений для наукових досліджень, який і використовується як базовий набір даних для виконання даної роботи.

У ньому представлено 6 різних типів дефектів, кожен із яких містить по 300 різних зображень. Таким чином, датасет складається з 1800 зображень. Типи поданих дефектів наступні (згідно з ГОСТ 21014-88): 1) crazing – волосовина, 2) inclusion – включення, 3) patches – плями, 4) pitted surface – горобина, 5) rolled-in-scale – вкатана окалина, 6) scratches – подряпини.

Якість вихідного набору даних досить висока і відмінно підходить для досліджень. Проте, під час навчання нейронної мережі важливо як якість вхідних даних, так і їх кількість. Оскільки 300 зображень на 1 дефект є невеликою кількістю, було вирішено провести аугментацію вхідного набору даних за допомогою нейромережевої бібліотеки Keras. Ця бібліотека має клас ImageDataGenerator, що є інтерфейсом для розширення навчальної множини даних за допомогою різних параметрів.

Основні параметри аугментації у класі ImageDataGenerator:

- 1) `samplewise_center` - встановлює середнє значення 0 для кожного прикладу у вибірці;
- 2) `samplewise_std_normalization` - ділить всі вхідні дані на середньоквадратичне відхилення;
- 3) `rotation_range` – ціле значення від 0 до 180, в діапазоні якого випадковим чином буде обраний кут для двовимірного повороту зображення;
- 4) `width_shift_range` – дробове значення, що не перевищує одиниці, яке відображає діапазон, в межах якого зображення буде зміщене щодо ширини відповідно випадковим чином;
- 5) `height_shift_range` – дробове значення, що не перевищує одиниці, яке відображає діапазон у межах якого зображення буде зміщене щодо висоти відповідно випадковим чином;
- 6) `rescale` - значення на яке будуть помножені (або поділені) всі дані перед застосуванням подальших обробок;
- 7) `shear_range` - інтенсивність зсуву (кут зсуву в радіанах) для випадкового застосування зсувних [8] перетворень (shearing transformations);
- 8) `zoom_range` – дробове значення, що відображає коефіцієнт для збільшення зображення в межах діапазону [`1 - zoom_range`; `1 + zoom_range`];
- 9) `horizontal_flip` – випадково застосовує горизонтальний поворот зображення;
- 10) `fill_mode` – задає спосіб заповнення елементів, які виходять за кордон, можливі варіанти: "constant", "nearest", "reflect" та "wrap".

Двома найбільш важливими для нейронної мережі параметрами є `samplewise_center` і `samplewise_std_normalization`, оскільки вони відповідають за нормалізацію даних, наводячи яскравість та розкид яскравості вхідних зображень в один діапазон.

Також важливий параметр `rescale` – необхідний для перетворення інтенсивності пікселів до діапазону від 0 до 1.

Крім, аугментації, необхідно було також розподілити дата сет на навчальну, валідаційну та тестову вибірки. Тому всі отримані зображення були розподілені на зазначені вибірки у відповідній пропорції – 75/15/15 % для кожного типу дефектів.

В роботі були реалізовані та адаптовані під завдання класифікації моделі на базі наступних архітектур штучних нейронних мереж: InceptionV3, MobileNet, MobileNetV2, а також модель простої згорткової мережі (або Base Convolution) для додаткового проведення порівнянь. Всім моделям надалі додається префіксне поєднання – Adapted For Defects Classifying (адаптована для класифікації дефектів), і кінцеві поєднання скорочуються до абревіатур, що формують назву моделей.

Для взаємодії із системою, відповідно до вимог, було реалізовано 2 інтерфейси: програмний (API) та графічний (у вигляді web-інтерфейсу).

IV. РЕЗУЛЬТАТИ

Для досягнення мети в рамках цієї роботи було виконано такі задачі: 1) виконано дослідження предметної галузі, визначено можливість використання нейронних мереж для вирішення задачі класифікації дефектів металопрокату; 2) побудовано модель нейронної мережі, що підходить для вирішення задачі класифікації видимих дефектів металопрокату; 3) виконано генерацію навчальної, тестової та валідаційної вибірок на основі бази даних дефектів поверхні гарячого металопрокату, здійснено аугментацію згенерованих вибірок; 4) спроектована та реалізована система класифікації дефектів.

V. ОБГОВОРЕННЯ

Основною метою роботи дослідження можливості використання нейромережових технологій для вирішення задачі класифікації дефектів металопрокату та створення програмної системи, що виконує класифікацію дефектів поверхні металопрокату з високою точністю (понад 90%) та високою швидкістю (не більше 1 секунди на 1 зображення) за рахунок використання штучних нейронних мереж.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальшими напрямками роботи будуть: реалізація можливості класифікації відразу кількох дефектів одного зображення та розробка програмного комплексу для виявлення та класифікації дефектів металопрокату у реальному часі.

VII. ВИСНОВКИ

В результаті роботи було досліджено можливість використання нейромережових технологій для вирішення задачі класифікації дефектів металопрокату та розроблено програмну систему для класифікації дефектів металопрокату на базі нейромережових технологій, здатну менш ніж за 1 секунду часу виконувати класифікацію дефекту на зображенні з точністю щонайменше 90 %.

ЛІТЕРАТУРА

- [1] Mazur I. *Quality Control system for a hot-rolled metal surface* // Frattura ed Integrità Strutturale, 2016. – Vol. 10. – No. 37. – P. 287-296.
- [2] Alkapov R.R. Automatic Visible Defect Detection and Classification System Prototype Development for Iron-and-Steel Works. / Alkapov R.R., Konyshov A.A., Vetoshkin N.A., Valkevich N.V., Kostenetskiy P.S. // *2018 Global Smart Industry Conference (GloSIC)*. Chelyabinsk, November 13-15, 2018, IEEE, 2018. – P. 1-8.
- [3] Batsuuri S., Ahn J., Ko J. *Steel surface defects detection and classification using SIFT and voting strategy*. // International Journal of Software Engineering and Its Applications, 2012. –Vol. 6. – P. 161-166.
- [4] Lemon S. *Avtomaticheskij osmotr poverhnosti polosy dlya obespecheniya sploshnogo kontrolya kachestva v linii*. // Chernye Metalli, 2003. – Vol. 4. – P. 49-54.
- [5] Stolzenberg M. *Sistemy osmotra poverhnosti listovoy staly: Rezultaty issledovaniy*. / Stolzenberg M., Gruber K., Henkenmayer H., Joncourt K. // Chernye Metalli, 2002. – Vol. 12. – P. 40-48.
- [6] Neogi N., Mohanta D., Dutta P. *Review of vision-based steel surface inspection systems*. // EURASIP Journal on Image and Video Processing, 2014. – Vol. 2014. – No. 1. – P. 1-50.
- [7] Guifang W. Design of online surface inspection system of hot rolled strips. / Guifang W., Kwak H., Jang S., Xu K., Xu J. // *2008 IEEE International Conference on Automation and Logistics*, 2008. – P. 2291-2295.
- [8] Vetoshkin N., Valkevich N., Kostenetskiy P. *Razrabotka polzovatel'skogo interfeysa dlya sistemy avtomaticheskogo kontrolya kachestva poverhnosti listovoy staly s pokrytiem*. // Parallel Computational Technologies (PCT'2018), 2018. – P. 400-402.

- [9] Caleb P., Steuer M. Classification of surface defects on hot rolled steel using adaptive learning methods. // *Proceedings of the Fourth International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies*, August 2000, Brighton, UK. IEEE, 2002. – Vol. 1. – P. 103–108.
- [10] Popat M., Barai S. Defect detection and classification using machine learning classifier. // *Proceeding of the 16th WCNDT 2004 - World Conference on NDT*, 2004. – P. 788-797.

Дослідження методів на основі нейронних мереж для аналізу тональності корпусу текстів

Катерина Островська, Денис Печений

*Кафедра інформаційних технологій і систем
Український державний університет науки та технологій
Дніпро, Україна*

Анотація — Об'єктом дослідження є методи з урахуванням нейронних мереж для аналізу тональності корпусу текстів. Для досягнення поставленої в роботі мети необхідно вирішити такі завдання: дивчитися теоретичний матеріал для навчання глибоких нейронних мереж та їх особливості стосовно обробки природної мови; вивчити документацію бібліотеки Tensorflow; розробити моделі згорткової та рекурентної нейронних мереж; розробити реалізацію лінійних та нелінійних методів класифікації на моделях мішка слів та Word2Vec; порівняти точність та інші показники якості реалізованих нейромережових моделей із класичними методами. Для візуалізації навчання використовується Tensorboard. У роботі показано перевагу класифікаторів на основі глибоких нейронних мереж над класичними методами класифікації, навіть якщо для векторних уявлень слів використовується модель Word2Vec. Найвищу точність для даного корпусу текстів має модель рекурентної нейронної мережі з LSTM-блоками.

Ключові слова — Штучні нейронні мережі, Глибокі нейронні мережі, Навчання з учителем, Глибоке навчання, Рекурентна нейронна мережа, LSTM, Згорткова нейронна мережа, Аналіз тональності тексту, Мішок слів, Word2vec.

I. ВСТУП

Щодня в Інтернеті з'являється величезна кількість контенту: користувачі висловлюють свою думку про фільми та події, залишають відгуки про різні продукти та послуги. Для вирішення завдань, пов'язаних із аналізом емоційного забарвлення тексту, використовуються методи аналізу тональності тексту.

Завдання автоматичного аналізу тональності тексту є досить популярним [1]. Найчастіше користувачі при виборі чогось (наприклад, до якого університету вступити) керуються думками інших людей. Тому набір опрацьованих думок становить значний інтерес для соціологів, маркетологів та власників бізнесу. Також система для аналізу тональності тексту може стати в нагоді на сайтах, де люди залишають відгуки: наприклад, користувач може позитивно оцінити фільм, випадково поставивши негативну оцінку - система автоматичного аналізу тональності тексту виправить помилку. Тому аналіз тональності текстів є важливим та актуальним завданням.

Автоматичний аналіз тональності тексту зазвичай застосовується на корпусах текстів, які містять рецензії чи відгуки. Також його можна застосувати для даних з блогів та соціальних мереж, щоб отримати громадську думку з того чи іншого питання або товару.

Визначення емоційної оцінки авторів рецензій перестав бути очевидним завданням. Стандартні методи класифікації тексту ізолюють слова за ознаками та застосовують різні методи вибору ознак, щоб знайти найбільш "важливе слово" у тексті: наприклад, речення "Мені подобається бігати" і "Мені не подобається бігати" вважатимуться однаковими. У ситуації, коли рецензія представлена лише позитивно забарвленим набором слів, а насправді негативною, стандартні методи класифікації тексту перестають бути ефективними. Методи машинного навчання дозволяють алгоритмам "розуміти" структуру речення та його семантичну структуру. Пропозиція буде представлена у вигляді вектора, в якому збережена структура речення та те, як слова пов'язані один з одним.

Для проведення чисельних експериментів були використані рецензії сайту Rotten Tomatoes [2] — набір із 5331 позитивних та 5331 негативних рецензій.

Потрібно побудувати бінарний класифікатор, визначальною, позитивною чи негативною виявилася рецензія. Як методи розглядаються згортова нейронна мережа і рекурентна нейронна мережа з LSTM-блоками, наївний класифікатор байесовського, метод опорних векторів і логістична регресія. Також було використано два варіанти векторних моделей подання тексту: мішок слів (Bag of Words) та Word2Vec.

В результаті навчання були отримані моделі нейронних мереж, що дозволяють з високою точністю визначати тональність рецензій, а також порівняння ефективності використання реалізованих методів.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Завданням, поставленою у роботі, є реалізація різних алгоритмів глибинного навчання для аналізу тональності тексту та порівняння їх ефективності з класичними (лінійними та нелінійними) алгоритмами класифікації тексту.

У дослідженні [3] представлені результати ефективності застосування різних архітектур згорткових нейронних мереж при використанні попередньо навченої моделі векторного уявлення word2vec.

Результати застосування нейронних мереж з LSTM-блоками та їх модифікаціями були представлені в дослідженні [4], а також було показано, що ця архітектура є найбільш ефективною для побудови бінарного та багатокласового класифікаторів для аналізу тональності текстів.

Також у роботі [5] стверджується, що можна значно покращити ефективність класифікатора, використовуючи попередньо навчені моделі векторних уявлень слів (наприклад, Word2Vec).

У цій роботі для побудови бінарного класифікатора будуть використовуватися рецензії, що складаються з коротких речень.

В даний час завдання побудови класифікатора (бінарного або багатокласового) для визначення тональності тексту вирішується за допомогою нейромережових моделей, оскільки ефективність архітектур, використаних у згаданих роботах, значно вища, ніж у класичних лінійних алгоритмів.

Існує безліч бібліотек для реалізації алгоритмів машинного навчання. Існують два типи фреймворків: символічні та імперативні. У символічних фреймворках набагато більше можливостей використовувати пам'ять багаторазово, а оптимізація на основі граф залежностей здійснюється автоматично. Найпопулярнішими символічними (symbolic) фреймворками нині є TensorFlow та Theano.

На відміну від Theano, Tensorflow не орієнтований лише на навчання нейронних мереж, тому можна використовувати колекції графів та черги як складові для високорівневих компонентів.

Якщо необхідно навчати масштабні моделі і використовувати багато зовнішньої пам'яті, Theano буде дуже повільно працювати через необхідність компіляції коду C/CUDA в бінарний код.

TensorFlow має прозору модульну архітектуру з безліччю фронт-ендів. В архітектурі Theano розібратися досить непросто: весь код – це Python, де код C/CUDA упакований як рядок Python. У такому коді складно орієнтуватися, його непросто налагоджувати та проводити рефакторинг. Більше того, візуалізація графів у TensorFlow реалізована значно ефективніше, ніж у Theano [6].

Векторні уявлення слів для лінійних алгоритмів будуть представлені двома моделями: Word2Vec та мішок слів (bag of words), а для класифікаторів на основі нейронних мереж буде використано лише модель Word2Vec. За допомогою інструменту для побудови векторних моделей Gensim буде навчено модель Word2Vec. З використанням TensorFlow будуть реалізовані згортова нейронна мережа та рекурентна нейронна мережа з LSTM-блоками.

III. МЕТОДИ

Процес обробки даних складається з наступних кроків:

- видалити розмітку HTML;

- видалити всі символи, окрім літер та пробілів;
- з отриманого набору слів видалити стоп-слова.

Наступним завданням є перетворення кожної рецензії на векторне уявлення. Для оцінки ефективності кожного з методів буде використано дві моделі векторного подання слів: мішок слів та Word2Vec.

У роботі логістична регресія використовується для передбачення ймовірності виникнення певної події.

Наївний класифікатор Байєса - простий імовірнісний класифікатор, заснований на застосуванні Теорема Байєса зі суворими припущеннями про незалежність елементів вектора ознак. Перевагою наївного класифікатора Байєса є мала кількість даних для навчання, необхідних для оцінки параметрів, необхідних для класифікації.

Випадковий ліс (random forest) – алгоритм полягає у використанні комітету вирішальних дерев. Класифікація об'єктів проводиться шляхом голосування: кожне дерево комітету відносить об'єкт, що класифікується, до одного з класів, і перемагає клас, за який проголосувало найбільшу кількість дерев. Оптимальна кількість дерев підбирається таким чином, щоб мінімізувати помилку класифікатора на тестовій вибірці. Використовує усереднення для підвищення точності прогнозування та контролю надлишкового припасування. Розмір підвиборки завжди збігається з розміром оригінальної вибірки.

Метод опорних векторів (SVM) - Основна ідея методу - переведення вихідних векторів у простір більш високої розмірності та пошук роздільної гіперплощини з максимальним зазором у цьому просторі. Стандартна функція scikitlearn (sklearn.svm.SVC) не підходить, оскільки тимчасова складність даного алгоритму квадратична. Ефективність методу опорних векторів значно знижується, якщо кількість ознак дуже велика. Він має велику гнучкість у виборі штрафів та функцій втрат і має краще масштабуватись для великої кількості зразків. Оптимізація функції втрат SVM за допомогою градієнтного спуску:

$$L(\omega, D) = \frac{1}{2} \|\omega\|_2^2 + C \sum_{i=1}^N \max(0, 1 - y_i (\omega^T x_i + b)), \quad (1)$$

де

$$D = \{(x_i, y_i)\}_{i=1}^N, x_i \in Rd \text{ and } y_i \in \{-1, +1\}. \quad (2)$$

Оптимізація функції втрат - (regularization_loss + hinge_loss). Для випадку word2vec буде використано SGDClassifier зі sklearn.linear_model з помилкою l2.

Згорткова нейронна мережа - Векторні подання даних здійснено за допомогою моделі Word2Vec (Skip-gram Model).

Як функція оптимізації використовується Adam (Adaptive Moment Estimation). Його відмінними рисами є:

- оцінка першого моменту обчислюється як ковзне середнє;
- оскільки оцінки першого та другого моментів ініціалізуються нулями, використовується невелика корекція, щоб результуючі оцінки не були зміщені до нуля.

Рекурентна нейронна мережа з LSTM-блоками - векторні подання даних здійснено за допомогою моделі Word2Vec (Skip-gram Model). LSTM вважаються найкращою архітектурою для аналізу тональності тексту. Нейронні мережі, складені з LSTM-модулів, особливо добре обробляють заперечення (negation), якщо у осередку є projection unit (тобто більше пам'яті мережі).

IV. РЕЗУЛЬТАТИ

Для класичних алгоритмів класифікації використання мішка слів як модель векторного подання слів точність передбачень не перевищувала 75.4% (логістична регресія), мінімальна точність становила 69.9% (лінійний метод опорних векторів).

Завдяки використанню моделі Word2Vec вдалося покращити точність передбачень майже всім методів (крім наївного байєсівського класифікатора): наприклад, точність лінійного методу опорних векторів стала 74.2%. Проте найефективнішим бінарним

класифікатором виявилася логістична регресія: її точність із використанням моделі Word2Vec дорівнює 76,6%. За допомогою згорткових нейронних мереж з моделлю Word2Vec вдалося отримати точність 79.9%.

Найефективнішою архітектурою для аналізу тональності тексту виявилася рекурентна нейронна мережа з LSTM-блоками. Її максимальна точність становила 83.0%.

Отримані експериментальні дані показують вищу ефективність роботи глибоких нейронних мереж проти класичними алгоритмами для аналізу тональності тексту.

V. ОБГОВОРЕННЯ

Результати дослідження свідчать, що використання глибоких нейронних мереж значно покращує точність аналізу тональності тексту.

Перевага рекурентної мережі на основі LSTM над згортковою нейронною мережею в галузі аналізу тональності вже була доведена в різних дослідженнях, проте важливо відзначити, що в даній роботі були реалізовані найпростіші архітектури глибоких нейронних мереж. Поліпшення параметрів моделі, використання більш розширеної моделі векторного уявлення слів Word2Vec, застосування attention-механізмів дозволить значно збільшити ефективність бінарного класифікатора для аналізу тональності на основі глибоких нейронних мереж.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Можливим напрямом для подальшої роботи може бути розпізнавання ключових слів, які роблять найбільший внесок у позитивний або негативний відгук. Введення модуля в експлуатацію.

VII. ВИСНОВКИ

Відповідно до поставленої мети — розробка алгоритмів глибокого навчання для аналізу тональності тексту та порівняння їх ефективності з іншими класифікаторами на основі алгоритмів машинного навчання — було реалізовано архітектуру згорткової нейронної мережі, рекурентної нейронної мережі з LSTM-блоками, а також проведено порівняння показників якості їх класифікації з іншими класифікаторами.

При використанні моделі мішка слів точність різних методів була значно вищою за випадкову (близько 70%), проте застосовуючи модель Word2Vec, вдалося значно покращити точність роботи алгоритмів (на кілька одиниць). Однак нейронні мережі показали найкращі результати.

Точність класифікатора на основі згорткової нейронної мережі виявилася 79.9%. Найвищу точність показав класифікатор на основі рекурентної мережі з LSTM-блоками — 83.3%.

ЛІТЕРАТУРА

- [1] B. Pang and L. Lee. *Opinion mining and sentiment analysis*. Foundations and Trends in Information Retrieval archive, 2008.
- [2] Данные рецензий, используемых в работе, sentence polarity dataset v1.0. <http://www.cs.cornell.edu/people/pabo/movie-review-data/>
- [3] Y. Kim. *Convolutional neural networks for sentence classification*. arXiv:1408.5882 [cs.CL], 2014.
- [4] K.S. Tai et al. *Improved semantic representations from tree-structured long short-term memory network*. arXiv:1503.00075 [cs.CL], 2015.
- [5] Q. Le and T. Mikolov. *Distributed representations of sentences and documents*. arXiv:1405.4053 [cs.CL], 2014.
- [6] K. Tran et al. Evaluation of deep learning toolkits. <https://github.com/zer0n/deepframeworks/blob/master/README.md>

Дослідження ефективності методів обробки зображень у системах розпізнання облич

Катерина Островська, Анастасія Петренко

*Кафедра інформаційних технологій і систем
Український державний університет науки та технологій
Дніпро, Україна*

Анотація — У роботі наведено різні підходи до класифікації методів розпізнавання осіб, наведено огляд та короткий опис відповідних методів. Крім того розглядаються існуючі підходи до побудови вибіркової сукупності, які можуть бути використані для формування навчального та тестового набору на основі наявних вихідних даних. Докладніше розглядаються вибрані для дослідження метод головних компонент і метод локальних бінарних шаблонів для вирішення завдання розпізнавання зображень. Описано принцип їхньої дії та алгоритм роботи. Описуються результати розпізнавання зображень осіб на зображеннях баз даних ORL та FERET, вибраними методами із застосуванням різних параметрів. Результати порівнюються між собою з метою визначення найефективніших варіантів. Розглядається вплив використання різних параметрів методу на ефективність розпізнавання, а також вплив ступеня нормалізації вихідного матеріалу на цю ефективність.

Ключові слова — *Метод Головних Компонент, Локальний Бінарний Шаблон, Лінійний Дискримінантний Аналіз, Нейронна Мережа, Згортована Нейронна Мережа, Багатошарова Нейронна Мережа, Приховані Марківські Моделі, Open Source Computer Vision Library.*

I. ВСТУП

Завдання виділення особи людини у природній чи штучній обстановці та наступної ідентифікації завжди перебувала у ряді найпріоритетніших завдань для дослідників, які працюють у галузі систем машинного зору та штучного інтелекту.

У багатьох дослідженнях, проведених у межах розв'язання цього завдання протягом кількох десятиліть, отримані найрізноманітніші результати, як у відмінних друг від друга методик, так методик, що використовують одні й самі алгоритми, так і для систем, заснованих на одному й тому ж алгоритмі.

Для об'єктивного порівняння різних методів розпізнавання зображень необхідно використовувати єдину методику перевірки та аналогічну вибірку зображень. Крім того, сам спосіб поділу вибірки на тестову і навчальну частини повинен мати мінімальну залежність від набору даних, що пред'являється.

Для того, щоб виділити найбільш ефективний метод ідентифікації зображення обличчя, необхідно застосувати математично обґрунтований метод оцінки того, наскільки успішні досліджувані методи на незалежному наборі даних. Крім того, слід оцінити вплив кольорової нормалізації зображення та виключити вплив його геометричної нормалізації. А також розглянути залежність результатів від обсягу тестової вибірки для того, щоб виділити найефективніший метод ідентифікації, мінімізувати вплив зовнішніх параметрів на оцінку результатів.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Завдання розпізнавання людини зображення обличчя діляться на три великих класи: пошук у великих базах даних, контроль доступу і контроль фотографій в документах. Вони різняться за вимогами, що надаються до систем розпізнавання, і за способами вирішення, тому є окремими класами. У рамках цієї роботи, актуальним є перший клас завдань. Оскільки будь-який архів, не так важливо відеоархів це чи просто набір зображень, є насамперед базою даних.

Різні і вимоги до помилок першого і другого роду для таких класів [1]. У експериментальній частині роботи, основним питанням є факт успішного розпізнавання як

такого, тобто ситуація, коли система відносить об'єкт, що пред'являється, до того класу, якому він відповідає, у всіх інших ситуаціях результат випробування вважається невдачею.

Слід також відзначити відмінність понять верифікації та розпізнавання (ідентифікації) [2]. У задачі верифікації невідомий об'єкт заявляє, що він належить до певного відомого класу. Система підтверджує чи спростовує це. У процесі розпізнавання слід зарахувати об'єкт, що належить невідомому класу до одного з відомих або зробити висновок, що цей об'єкт не належить до жодного з відомих класів.

Часто у великих базах даних потрібно знайти такі зображення, які найбільше схожі на ті що подаються на вхід [3]. Пошук повинен проводитись за прийнятний період часу. Один із варіантів полягає в тому, що в базі даних зберігаються деякі заздалегідь вилучені набори важливих ознак, що визначають зображення із граничною однозначністю. У цій ситуації вимоги до точності не є настільки критичними як у завданнях контролю доступу або контролю документів. До цього класу, насамперед, слід зарахувати спосіб головних компонентів [2, 4]. Коефіцієнти, отримані при розкладанні вхідного зображення на головні компоненти, використовуються для порівняння зображень шляхом обчислення Евклідової відстані.

Якість, що є у методів розпізнавання, інваріантні до таких варіацій як зображення, масштаб, усунення, умови освітленості і т.д. Крім цього, є деякі способи обробки зображення з метою досягнення подібної інваріантності при використанні інших методів.

III. МЕТОДИ

Структурні підходи до розпізнавання об'єктів ґрунтуються на автоматичного підстроювання до властивостей оброблюваних даних, що дозволяє розробити систему, що розпізнає, з прийнятними характеристиками. До другої категорії належать методи, що характеризуються скороченням розмірності даних.

Складність завдання розпізнавання об'єктів полягає в тому, що характерні ознаки є багатовимірними векторами, що в сукупності описують деякі властивості об'єктів, що розглядаються, наприклад: текстури, набір моментів, гістограми і т. п. Серед адаптивних методів можна виділити такі різновиди:

- Метод опорних векторів;
- Штучні нейронні мережі;
- Приховані Марківські моделі;
- Метод гнучкого порівняння на графах.

Основною метою холістичних методів є знаходження оптимального зменшення розмірності множини ознак, що дозволяє уявити основні особливості кожного об'єкта. Для цього можна використати:

- лінійний дискримінантний аналіз;
- метод головних компонентів.

Суть кластерних підходів для порівняння осіб полягає в поділі зображення обличчя на деяку кількість ділянок, до кожної з цих ділянок можна застосувати холістичний підхід і порівнювати їх попарно незалежно від інших ділянок. У такому разі можна сподіватися на те, що якщо деякі пари ділянок порівнюваних зображень не збігатимуться між собою в силу вищезазначених заважаючих факторів на кшталт тіней і перекриттів, пари ділянок, що залишилися, забезпечать потрібний ступінь взаємної кореляції, якої буде достатньо для правильного підсумкового рішення неспівпадіння осіб.

Вихідний набір зображень був сформований з двох вільнорозповсюджуваних баз зображень осіб ORL (Olivetti - Olivetti Research Laboratory) [5] та FERET.

У основу ORL входять передні зображення осіб. База містить 400 передніх зображень 40 чоловік. Всі зображення в базі напівтонові з 256 градаціями яскравості. Розмір кожного зображення – 92×112 пікселів. Файли представлені у форматі PGM.

База зображень осіб FERET – це стандартний набір, який використовується для оцінки систем розпізнавання облич. Технологія розпізнавання осіб – FERET – програма під управлінням агентства передових оборонних дослідницьких проектів DARPA та
2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

Національного інституту стандартів та технологій США [6]. Розмір зображень облич бази даних FERET дорівнює 512×76844 пікселів.

Мета даної роботи – дослідження ефективності методів розпізнавання та ідентифікації осіб. Весь набір зображень був оброблений за допомогою каскадів Хаара, виявлені каскадом області зображень, виділені як зображення обличчя та стали основою подальшого дослідження. Таким чином, для обробки пропонувалися лише ті частини зображень, де машина виявила особу, і ефективність методів виявлення обличчя не вплинула на результати дослідження.

IV. РЕЗУЛЬТАТИ

У роботі розглядаються деякі аспекти використання алгоритмів ідентифікації особистості зображення обличчя, які можуть впливати на ефективність такого розпізнавання. Для досягнення більшої об'єктивності було вирішено такі питання:

1. Розроблено алгоритм формування тестової та навчальної вибірки з неоднорідного вихідного набору даних;
2. З використанням цього алгоритму проведено дослідження та порівняння ефективності методів ідентифікації особи за зображенням особи;
3. Виявлено особливості вихідного набору даних, при яких певний метод ідентифікації має великі переваги;

Після процедури виділення осіб у підсумкових наборах виявилось 515 класів, які відповідають обраним критеріям. 477 їх у базі FERET і 38 у базі ORL. Усього було виділено 4705 зображень осіб (4366 належали класам бази FERET, 339 – ORL). У кожному класі з бази ORL вийшло від 5 до 10 (3 – 8 у навчальній вибірці) зображень, а у класах з бази FERET від 4 до 36 (2–28 для навчальної вибірки).

4. Визначено, які вхідні параметри методів більшою мірою впливають на ефективність та точність розпізнавання методів.

5. На підставі результатів дослідження було сформульовано рекомендації щодо використання методів розпізнавання осіб.

V. ОБГОВОРЕННЯ

Порівняння результатів обробки дозволило зробити такі висновки:

1. В цілому, метод локальних бінарних шаблонів має більшу ефективність, ніж метод головних компонент як на даних бази FERET, так і на даних ORL.
2. Зі збільшенням кількості зображень у класі ефективність ідентифікації зростає на матеріалі обох баз для обох методів з будь-яким поєднанням вибраних параметрів;
3. На даних бази ORL зі збільшенням кількості зображень у класі різниця між ефективністю методів зменшуються.
4. На даних бази FERET зі збільшенням кількості зображень у класі різниця між ефективністю методів незначно збільшуються.
5. На даних бази FERET перевага в ефективності методу локальних бінарних шаблонів над методом головних компонентів більша, ніж на матеріалі бази ORL.
6. Вплив світлової нормалізації на результати методу локальних бінарних шаблонів мінімальний для обох баз.
7. Метод головних компонент дещо ефективніший при світловій нормалізації на матеріалі бази FERET, вплив цього параметра менш помітно ніж на матеріалі бази ORL.
8. На даних бази ORL збільшення радіусу знижує ефективність методу локальних бінарних шаблонів на FERET ситуація зворотна.
9. Найнижчий показник ефективності ідентифікації зображень осіб у методу головних компонентів без світлової нормалізації, при малій кількості зображень у класі.
10. Найкращі результати ефективності ідентифікації зображень облич показує метод локальних бінарних шаблонів з мінімальним радіусом без світлової нормалізації на матеріалі бази ORL при великій кількості зображень у класі.

Можна зробити висновок, що метод локальних бінарних шаблонів, практично у всіх досліджених варіантах поєднання вхідних параметрів і вихідних даних, представлених для

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

обробки, перевищує ефективність розпізнавання методом основних компонентів. Ця перевага стає очевидною, коли вихідний набір даних має досить великий обсяг, тим самим знижується ймовірність помилки в результатах дослідження, нав'язаної специфікою матеріалу, що обробляється та фактором випадковості. При використанні в якості вихідного матеріалу зображень спочатку тих, що мають меншу «ступінь нормалізації» така перевага була куди більш відчутна. З цієї причини, в реальних умовах, можна рекомендувати використання методу локальних бінарних шаблонів.

У процесі дослідження стало очевидним, що навчання класифікаторів при великому обсязі вихідних даних дуже вимогливе до апаратних ресурсів і часу. Частково цю проблему вдалося вирішити шляхом розпаралелювання процесу обробки даних і запуску їх у кілька потоків. Однак, на окремій робочій станції обмеження продуктивності все ж таки виявилися дуже відчутними. Внаслідок цього для вирішення масштабних завдань рекомендується використовувати розподілені обчислення безпосередньо для обробки даних та продуктивну систему зберігання даних для запису та зберігання навчених класифікаторів.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальшими напрямками роботи буде: реалізація та розробка програмного комплексу для дослідження ефективності методів обробки зображень у системах розізнання осіб у реальному часі.

VII. ВИСНОВКИ

В ході роботи було досліджено різні підходи до класифікації методів розпізнавання осіб, проаналізовано відповідні методи. Розглядаються існуючі підходи щодо побудови вибіркової сукупності.

За результатами вивчення цих питань були обрані методи для дослідження та порівняння ефективності ідентифікації людини за зображенням особи: метод головних компонентів та метод локальних бінарних шаблонів.

Вихідним матеріалом для оцінки та порівняння ефективності методів стали бази даних ORL та FERET. Для об'єктивного порівняння результатів ідентифікації було підготовлено кілька варіантів вибірок зображень з розбиттям наявних наборів зображень у класах на різну кількість елементів. Шляхом перехресної перевірки результатів для аналогічних наборів даних, була проведена обробка отриманих наборів за допомогою досліджуваних методів із застосуванням різних параметрів.

ЛІТЕРАТУРА

- [1] Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2018). *Faceforensics: A large-scale video dataset for forgery detection in human faces*. arXiv preprint arXiv:1803.09179.
- [2] Mashtalir, S., & Mikhnova, O. (2017). *Detecting significant changes in image sequences*. In *Multimedia Forensics and Security*, pp. 161-191.
- [3] Xi, C., Chen, J., Zhao, C., Pei, Q., & Liu, L. (2018, February). *Real-time hand tracking using kinect*. In *Proceedings of the 2nd International Conference on Digital Signal Processing*, pp. 37-42.
- [4] Naik, Ganesh R., ed. *Advances in Principal Component Analysis*. Singapore: Springer Singapore, 2018. // URL: <http://dx.doi.org/10.1007/978-981-10-6704-4>.
- [5] The ORL Database of Faces, /AT&T Laboratories Cambridge 2018. // URL: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- [6] The Color FERET Database, / The National Institute of Standards and Technology (NIST) 2009. // URL: <http://www.nist.gov/itl/iad/ig/colorferet.cfm>

MitM кібератаки на локально-керуючий рівень кіберфізичної системи розумних мереж енергозабезпечення: визначення доменної області та поведінка системи

Анастасія Нічепорук, Андрій Нічепорук, Андрій Казанцев, Олег Парніцький

*Кафедра комп'ютерної інженерії та інформаційних систем
Хмельницький національний університет
Хмельницький, Україна*

Анотація—В роботі було здійснено формалізацію MitM кібератаки (людина по середині) на локально-керуючий рівень кіберфізичної системи розумних мереж енергозабезпечення, визначено доменну область дослідження та формалізовано поведінку системи під впливом MitM кібератаки.

Ключові слова — *Кіберфізична система розумних мереж енергозабезпечення, MitM кібератаки, Поведінка системи, DNP3, Локально-керуючий рівень.*

I. ВСТУП

Кіберфізична система (КФС) розумних мереж енергозабезпечення є реалізацією загальної концепції відкритих кіберфізичних систем, у яких протікають інформаційні процеси пов'язані із наглядним управлінням фізичними пристроями та збереженням, перетворенням й обробкою даних від цих пристроїв, з метою автоматизації та оптимізації виробничих процесів.

З точки зору безпеки та захисту даних рівень наглядового управління (комунікаційна мережа та структурні об'єкти кіберфізичної системи) є найбільш захищеним до кібератак, що зумовлено їх наближенням до традиційних ІТ інфраструктур, для яких визначено та реалізовано застосування як апаратних так і програмних механізмів захисту та виявлення зловмисних впливів, що представлені зокрема брандмауерами, системами виявлення вторгнень, шифруванням каналів передачі даних, резервуванням, антивірусними засобами, тощо. В той самий час рівень локальний керуючий рівень є вразливим до кібератак, що пов'язано в першу чергу із застосуванням спеціальних протоколів передачі інформації DNP3, Modbus, IEC 61850 для яких показники захищеності та конфіденційності передачі інформації відійшли на другий план, поступившись місцем ефективності передачі інформації між фізичним та кібер середовищем. Тому кібератаки MitM на кіберфізичну систему розумної мережі енергозабезпечення досліджуватимуться в даній роботі саме для локального керуючого рівня.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Методи виявлення кібератак типу MitM у системах розумних мереж енергозабезпечення можна розділити на шість категорій: методи на основі локалізації, методи на основі залучення штучного інтелекту, моделі прогнозування, методи на основі характеристик каналу, методи на основі фільтрації, а також системи виявлення вторгнень. Значну частину методів складають підходи на основі залучення штучного інтелекту.

Категорія методів на основі штучного інтелекту складається з різних алгоритмів машинного та глибокого навчання, методів інтелектуального аналізу даних, еволюційних алгоритмів і методів нечіткої логіки. Наприклад, автори роботи [1] використовували машинне навчання для виявлення атак, а саме випадковий ліс, метод опорних векторів і нейронну мережу. Їх результати показують, що запропонована методика на основі випадкового лісу забезпечує досить високу точність. У [2] автори також використовували алгоритми машинного навчання для виявлення атак соціальної інженерії. Підхід використовує навчання без вчителя, що означає відсутність попередніх знань про спостережувані кібератаки. Вони порівнювали продуктивність різних алгоритмів машинного навчання (зокрема метод опорних векторів нейронні мережі) з точки зору надійності, точності та швидкості. Результати їх експериментів демонструють, що метод опорних векторів дають кращі результати порівняно з іншими алгоритмами. У [3] автори використовували методи машинного навчання для виявлення атак грубої сили на протокол Secure Shell. Автори використовували різні класифікатори, включаючи метод k-ближніх сусідів, дерева рішень та наївний байєсів класифікатор. Спільним недоліком даних методів не достатній рівень достовірності виявлення, що зумовлено тим, що в якості ознак для застосування методів машинного навчання були використанні атомарні ознаки. Тому перспективним напрямком є створення комплексних ознак на основі побудови параметризованих процедур кібератак та правил їх виявлення, які б лягли в основу системи виявлення вторгнень із залученням методів машинного навчання

III. ФОРМАЛІЗАЦІЯ MITM КІБЕРАТАКИ НА ЛОКАЛЬНО-КЕРУЮЧИЙ РІВЕНЬ КІБЕРФІЗИЧНОЇ СИСТЕМИ РОЗУМНИХ МЕРЕЖ ЕНЕРГОЗАБЕЗПЕЧЕННЯ

Кібератака MitM у КФС є пасивною чи активною формою втручання у комунікаційний канал обміну інформацією між пристроями системи, шляхом здійснення легітимізації своєї присутності, стаючи тим самим повноцінним учасником комунікаційного обміну [4, 5]. Таким чином, зловмисник, що реалізує атаку MitM на розумну мережу енергозабезпечення, може виконувати кібератаки на введення помилкових даних (false data injection) і введення помилкових команд (false command injection), які можуть поставити під загрозу роботу всієї енергосистеми, наприклад оцінки стану, її диспетчеризацію та автоматичне керування генерацією електричної енергії [6].

При розгляді кібератак на локально керуєний рівень кіберфізичної системи розумної мережі енергозабезпечення будемо вважати, що зловмисник, який реалізує кібератаку MitM є складовою частиною самої КФС, тобто розширює загальну кількість об'єктів у множині Ω :

$$\Omega = \{o_j\}_{j=1}^{N_\Omega} \cup o_{MitM} \quad (1)$$

де o_{MitM} – об'єкт кіберфізичної системи, що виконує компрометування каналу зв'язку та реалізує MitM кібератаки, Ω – повна група об'єктів КФС.

Це твердження виходить з огляду на те, що для початку проведення вектору атаки MitM на складову частину КФС (комунікаційний канал між MTU та віддаленим пристроєм телеметрії), зловмисник має бути представлений у вигляді об'єкта цієї системи, тобто бути учасником обміну інформацією.

Присутність в складі КФС об'єкту, що реалізує MitM кібератаки на об'єкти з множини $\Omega \setminus o_{MitM}$ призводить до того, що реалізується *параметризована процедура атаки*, результатом якої є надходження помилкових даних і/або команд на цільовий об'єкт кіберфізичної системи $o_{target} \in \Omega$. Результатом такої процедури для цільового об'єкта o_{target} є зміна його параметрів, наприклад параметр “стан” автоматичного комутаційного перемикача зміниться із “закритий” на “відкритий” або навпаки. Таким чином, змінюється стан самого цільового об'єкту, що призводить до формування нової траєкторії функціонування об'єкта (зміни спостережуваної траєкторії на нову). Траєкторією об'єкта КФС назвемо не порожню кінцеву

послідовність зміни стану. Визначимо функцію переходу ξ , що визначає перехід об'єкта з стану s_k у стан s_{k+1} під впливом події r , в момент часу t_i наступним чином:

$$\xi : s_k \xrightarrow{r, t_i} s_{k+1} \quad (2)$$

З урахуванням траєкторій функціонування об'єктів подамо узагальнене представлення кібератаки MitM наступним рисунком (рис.1):

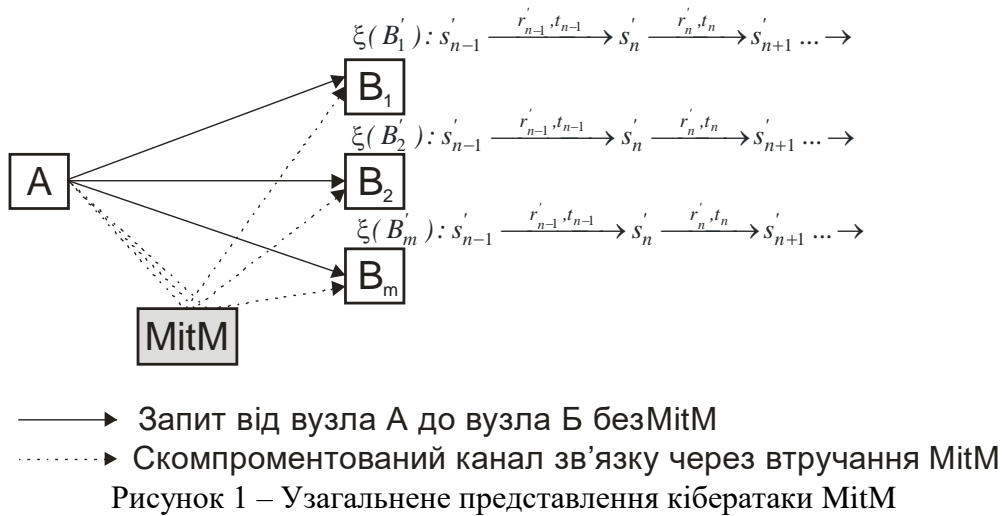


Рисунок 1 – Узагальнене представлення кібератаки MitM

При відсутності компрометації каналу зв'язку між вузлом A (головна станція) та множиною віддалених станцій B_1, B_2, \dots, B_m (підпорядковані вузли) траєкторії об'єктів КФС будуть мати вигляд:

$$X_c = \begin{cases} \xi(B_1) : s_{n-2} \xrightarrow{r_{n-2}, t_{n-2}} s_{n-1} \xrightarrow{r_{n-1}, t_{n-1}} s_n \xrightarrow{r_n, t_n} s_{n+1} \dots \rightarrow \\ \xi(B_2) : s_{n-2} \xrightarrow{r_{n-2}, t_{n-2}} s_{n-1} \xrightarrow{r_{n-1}, t_{n-1}} s_n \xrightarrow{r_n, t_n} s_{n+1} \dots \rightarrow \\ \dots \\ \xi(B_m) : s_{n-2} \xrightarrow{r_{n-2}, t_{n-2}} s_{n-1} \xrightarrow{r_{n-1}, t_{n-1}} s_n \xrightarrow{r_n, t_n} s_{n+1} \dots \rightarrow \end{cases} \quad (3)$$

де X_c – поведінка системи без компрометації каналу обміну інформації

Тоді після проведення кібератаки MitM та компрометації каналів зв'язку між A та B_1, B_2, \dots, B_m можна отримати наступні траєкторії:

$$X_{MitM} = \begin{cases} \xi(B'_1) : s'_{n-2} \xrightarrow{r'_{n-2}, t'_{n-2}} s'_{n-1} \xrightarrow{r'_{n-1}, t'_{n-1}} s'_n \xrightarrow{r'_n, t'_n} s'_{n+1} \dots \rightarrow \\ \xi(B'_2) : s'_{n-2} \xrightarrow{r'_{n-2}, t'_{n-2}} s'_{n-1} \xrightarrow{r'_{n-1}, t'_{n-1}} s'_n \xrightarrow{r'_n, t'_n} s'_{n+1} \dots \rightarrow \\ \dots \\ \xi(B'_m) : s'_{n-2} \xrightarrow{r'_{n-2}, t'_{n-2}} s'_{n-1} \xrightarrow{r'_{n-1}, t'_{n-1}} s'_n \xrightarrow{r'_n, t'_n} s'_{n+1} \dots \rightarrow \end{cases} \quad (4)$$

де X_{MitM} поведінка системи при компрометації каналу обміну інформації

Таким чином починаючи від моменту часу t_{n-1} стан об'єктів B_1, B_2, \dots, B_m змінюється, що призводить до зміни їх траєкторій функціонування. В результаті для об'єктів B_1, B_2, \dots, B_m зміниться їх поведінка, яка включатиме окрім траєкторій $\xi(B_1)$, $\xi(B_2)$ та $\xi(B_m)$ також й траєкторії $\xi(B'_1)$, $\xi(B'_2)$ та $\xi(B'_m)$.

Якщо порівняти вирази 3 та 4, і припустити, що стани s об'єктів є невідомі, то з точки зору процесу виявлення кібератак типу MitM ключовим моментом є визначення множини подій $r_i \in R$, що призводять до зміни стану цільового об'єкта o_{target} та формують нові траєкторії. В такому випадку, якщо розглядати комунікаційний канал між MTU та віддаленим пристроєм телеметрії, подіями виступатимуть команди, що надходять від головного вузла до підлеглих (віддалені станції з підключеним набором давачів та виконавчих механізмів), а також дані,

що надходять на головну станцію у відповідь на отримання керуючих команд. Ці дані та команди передаються по каналах комунікації через множину TCP/IP пакетів, у яких інкапсульовано DNP3 пакети, формуючи тим самим мережевий трафік.

IV. РЕЗУЛЬТАТИ

Для проведення експериментів було проведено моделювання MitM кібератак на локально керуючий рівень кіберфізичної системи. Даний рівень (домен) було змодельовано засобами openDNP3 бібліотеки [7], таким чином, що два хости в мережі обмінювались даними через протокол DNP3, що працював поверх стеку TCP/IP. В даній системі один хост представляв головний вузол (master, може представляти програмований логічний контролер), інший віддалений пристрій (outstation). Посередині між цими вузлами додано ще один хост, що виконував функції зловмисника. Результати проведеного дослідження дозволили сформуванню набір комплексних ознак зокрема: ознаку виявлення кібератаки на двійкові команди керування, аналогові команди керування, а також ознаку інвертування двійкових команд керування.

V. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В якості напрямку подальшого дослідження буде обрано дослідження відомих різновидів MitM кібератак, їх формалізацію, побудова параметризованих процедур кібератак та правил їх виявлення, які б лягли в основу системи виявлення вторгнень на локально-керуючий рівень кіберфізичної системи розумних мереж енергозабезпечення.

VI. ВИСНОВКИ

В роботі було здійснено формалізацію MitM кібератаки на локально-керуючий рівень кіберфізичної системи розумних мереж енергозабезпечення, визначено доменну область дослідження та формалізовано поведінку системи під впливом MitM кібератаки. З огляду на процес виявлення кібератак типу MitM важливим завданням є їх аналіз та формалізація параметризованих процедур їх впливу на мережевий трафік, що дозволить сформуванню сценарії впливу та виокремити набір ознак. Такі сценарії можуть бути використані для формування бази правил, відображення яких на набір ознак дозволить здійснити процес виявлення кібератак типу MitM.

ЛІТЕРАТУРА

- [1] W. Wang, and Z. Lu, "Cyber Security in the Smart Grid: Survey and Challenges", Computer networks, 2013, vol. 57, no. 5, pp.1344-1371.
- [2] C. Smutz, and A. Stavrou, "Malicious PDF Detection using Metadata and Structural Features", Annual computer security applications conference, 2012, pp. 239-248.
- [3] S. Fatima, and N. Kaabouch. "Social Engineering Attacks: A Survey." Future Internet, 2019, vol. 11, no. 4, p. 89.
- [4] J. Bai, S. Hariri and Y. Al-Nashif, "A Network Protection Framework for DNP3 over TCP/IP protocol," 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA), 2014, pp. 9-15, doi: 10.1109/AICCSA.2014.7073172.
- [5] Hao Li, Guangjie Liu, Weiwei Jiang and Yuewei Dai, "Designing snort rules to detect abnormal DNP3 network data," 2015 International Conference on Control, Automation and Information Sciences (ICCAIS), 2015, pp. 343-348, doi: 10.1109/ICCAIS.2015.7338690.
- [6] Ö. Sen et al., "Investigating Man-in-the-Middle-based False Data Injection in a Smart Grid Laboratory Environment," 2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), 2021, pp. 01-06, doi: 10.1109/ISGTEurope52324.2021.9640002.
- [7] URL: <https://dnp3.github.io/>

Постановка задачі розробки програмного забезпечення для оцінювання лісистості та оброблення інформації про лісові насадження

Олександр Мельников, Вікторія Денисенко

*Кафедра інтелектуальних систем прийняття рішень
Донбаська державна машинобудівна академія
Краматорськ, Україна*

Анотація—Розглянуто проблему збереження лісових насаджень та використання для цього спеціалізованого програмного забезпечення. Описано створений раніше додаток, який дозволяє вирішувати, чи перевищує кількість гектарів, де проходить вирубка, певного значення, на прикладі Придонецького лісництва Ізюмського району Харківської області. Сформульовано задачу розробки спеціалізованого програмного забезпечення – системи підтримки прийняття рішень для оцінювання лісистості, оброблення інформації про лісові насадження та виявлення незаконної вирубки лісу. Наведено прототип реалізації деяких пунктів задачі.

Ключові слова—*Збереження лісу, Програмне забезпечення, Система підтримки прийняття рішень.*

I. ВСТУП

Ліс – це сукупність землі, рослинності, в якій переважають дерева та чагарники, тварини, мікроорганізми та інші природні складові, що у своєму розвитку біологічно взаємопов'язані, впливають одна на одну і на довкілля. Знищення таких екосистем здатне призвести до найсерйозніших наслідків. В Україні питання збереження лісів регулюється Конвенцією про біологічне різноманіття та Конвенцією про охорону дикої флори та фауни та природних середовищ існування у Європі (Бернська конвенція) та ще низкою законодавчих актів [1-2]. Створення та використання спеціалізованого програмного забезпечення для оцінювання лісистості, оброблення інформації про лісові насадження та виявлення незаконної вирубки лісу може здійснити суттєвий вплив щодо покращення ситуації у цьому напрямку.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Є низка сервісів, що використовують відкриті дані у сфері лісового господарства, і в зручній формі надають необхідну інформацію користувачам [3]. Через онлайн-ресурс і мобільний застосунок «Екосфера» будь-хто може подати скаргу на підозрілі вирубування. Для цього потрібно позначити місце на Google-карті та завантажити фото- або відеоматеріали з місця події. Аналітичний портал та інтерактивна карта вирубувань ДП «Лісогосподарський інноваційно-аналітичний центр» [4] допоможуть перевірити наявність дозвільних документів на заготівлю деревини та перевірити законність здійснення вирубувань. Якщо на карті лісорубного квитка немає, а вирубування в лісі триває, то необхідно звернутися до лісництва чи лісгоспу, на території якого це відбувається, та повідомити про вирубку, якої немає на карті. Але не завжди ці додатки можуть знайти такі невеличкі лісництва, як, наприклад, Придонецьке. Автори сформулювали [5] та розв'язали [6] задачу створення програмного забезпечення для виявлення незаконної вирубки лісу на прикладі селища Співаківка в Ізюмському районі Харківської області [7]. Розроблений додаток містить дані про номер гектара, стан вирубки на ньому та невеличку інформацію, щоб користувач мав можливість дізнатися про лісові насадження, які є на даній території; стан вирубки; бригаду, яка проводить вирубку на даному гектарі; дату останньої вирубки;

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

дату запланованої вирубки. На рис. 1 наведено приклад вікна з елементами створеної бази даних. Система приймає рішення, чи перевищується відсоток вирубки в даному лісництві (норма для вирубки дорівнює 34% на лісництво) – рис. 2.

Гектар	Вид насаджень	Стан вирубки	Дата останньої вирубки	Дата запланованої вирубки	Бригада, яка вирубує
333	Сосна	-	13.02.2016	-	-
334	Сосна	-	12.01.2000	19.10.2022	-
335	Сосна	-	-	15.01.2022	-
336	Сосна	-	29.09.2011	23.10.2029	-
337	Сосна	-	12.11.2015	-	-
338	Сосна	+	14.08.2017	-	Придонецький лісгосп N
339	Сосна	-	-	02.05.2022	-
340	Сосна	+	-	12.12.2021	Ізюмський лісгосп №6
341	Сосна	+	11.09.2014	-	Придонецький лісгосп N
342	Сосна	+	-	-	Ізюмський лісгосп №2
375	Сосна	-	13.04.2006	12.09.2025	-
376	Сосна	-	21.09.2020	-	-
377	Сосна	-	11.03.2006	14.08.2029	-
378	Сосна	-	10.04.2021	-	-
379	Сосна	-	-	2.09.2022	-
380	Сосна	-	15.05.2000	04.07.2025	-
381	Сосна	-	-	1.10.2023	-
382	Сосна	-	2.04.2020	-	-

Рисунок 1 – Приклад вікна з елементами створеної бази даних

Рисунок 2 – Розрахунки системи

III. МЕТОДИ

Таким чином, наразі є додаток, який містить карту та базу даних по конкретному лісництву (номер ділянки, що переважає вид рослинності; інформація про вирубку), причому обробка бази вже реалізована. Далі для створення спеціалізованого програмного забезпечення – системи підтримки прийняття рішень для оцінювання лісистості та іншого оброблення інформації про лісові насадження заплановані наступні кроки:

1. Створити модуль обробки гугл-карт та аналізу отриманих даних для визначення ступеня «лісистості» кожної ділянки (гектара, квадрата...) та порівняння з минулими роками.
2. Об'єднати новий модуль з чинною базою даних.
3. Провести аналіз наявних моделей і практичних рішень.
4. Розробити математичні моделі та застосувати сучасні методи.

По першим двом пунктах передбачається наступне. У базу даних додаються нові поля: GPS-координати кожної ділянки (x_1 , y_1 , x_2 , y_2) і рівень «лісистості» (розраховується). Система при виборі відповідного пункту меню починає великий цикл по всій базі даних і виконує наступне:

- завантажує та показує на екрані гугл-карту ділянки (по наявних координатах «вершин прямокутника»);
- виконує попередню обробку (очевидно, потрібно зробити інверсію кольорів тощо);
- зберігає малюнок у теці року з ім'ям файлу, що містить номер ділянки (наприклад, у теку 2022 зберігаємо файли Image001.bmp, Image002.bmp, ..., Image333.bmp, ...);
- виконує розрахунок (найпростіший – насиченість зеленими кольорами, та інші);
- зберігає показник у базі даних.

Можна відсортувати базу даних по зростанню або спаданню «показника лісистості», щоб було видно, на яких ділянках можна починати вирубку, а на яких, навпаки, починати висаджувати новий ліс.

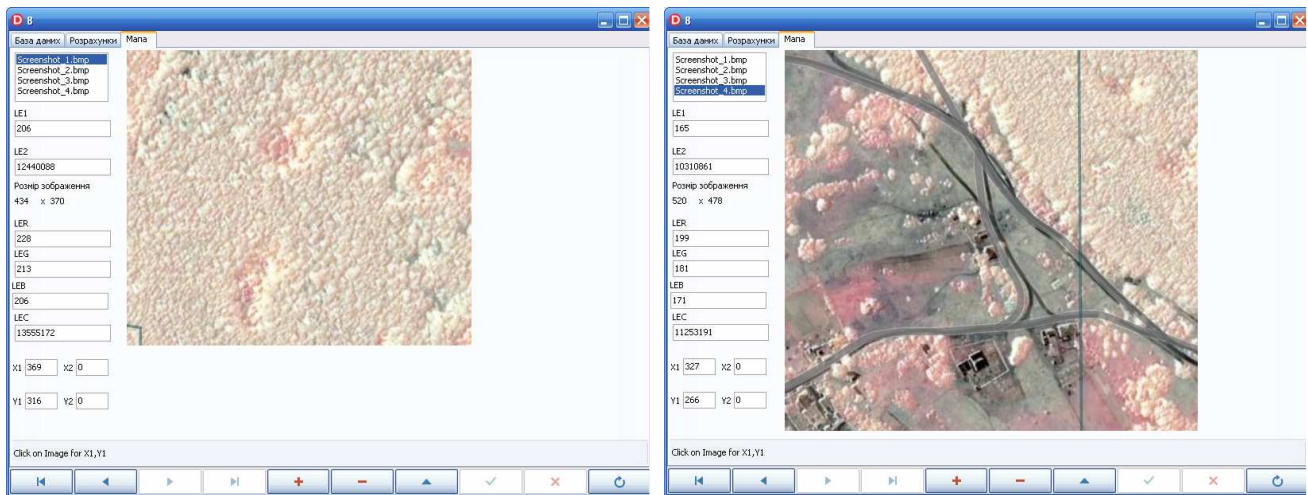
Коли будуть в наявності дані за кілька років, виконується нове завдання: проводиться порівняння насиченості зеленими кольорами, і робимо висновок, наскільки там зменшилася лісистість. У перспективі (якщо дані будуть більше, ніж за три роки) буде зроблено прогнозування зміни лісистості на кожній ділянці. (Файли із зображеннями зі супутника треба зберігати в одну пору року).

Головне завдання – одержати координати кожної ділянки. Це можна зробити двома шляхами:

- «вручну», завантаживши гугл-карту лісництва (у нашому випадку – Придонецького), і, акуратно вказуючи мишею на границі ділянок, зберегти всі координати у файлі (або відразу в базі даних);
- зробивши окремих модуль: у додатку з'являється гугл-карта району (Придонецького лісництва), потім завантажується наявне фото (схема), і програма сама робить розрахунок ділянок.

IV. РЕЗУЛЬТАТИ

Було створено перший прототип нового модуля додатка: на рис 3 наведено порівняння двох ділянок з селища Придонецьке: на першому (а) є в наявності тільки лісові насадження, а на другому (б) – ще й господарські будови. Додаток правильно розрахував, що насиченість зеленим кольором першої ділянки в 1,25 раза вище. Приклад також демонструє, що ділянки можуть бути різного розміру.



а

б

Рисунок 3 – Порівняння ділянок

V. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Далі треба здійснити програмну реалізацію усіх сформульованих завдань.

VI. ВИСНОВКИ

Було сформульовано задачу створення спеціалізованого програмного забезпечення для оцінювання лісистості, оброблення інформації про лісові насадження та виявлення незаконної вирубки лісу, яке може здійснити суттєвий вплив щодо покращення ситуації у цьому напрямку.

ЛІТЕРАТУРА

- [1] Основні поняття про ліси та чому важливе їх збереження. [Електронний ресурс]. Режим доступу: <https://mcl.kiev.ua/osnovnyue-ponjatija-o-lesah-i-rochemu-vazhno-ih-sohranenie/>. Дата звернення: Вер. 15, 2022.
- [2] Знищення лісів: причини і наслідки. [Електронний ресурс]. Режим доступу: <https://tvir.biographiya.com/znishhennya-lisiv-prichini-i-naslidki/>. Дата звернення: Вер. 15, 2022.
- [3] Вирубвання лісів України – Всеукраїнська екологічна ліга. [Електронний ресурс]. Режим доступу: <https://www.ecoleague.net/forumy-konferentsii-kruhli-stoly-seminary/ekolohichni-viiny/item/572-vyrubuvannia-lisiv-ukrainy>. Дата звернення: Вер. 15, 2022.
- [4] Українське державне проектне лісовпорядне виробниче об'єднання. [Електронний ресурс]. Режим доступу: <https://www.lisproekt.gov.ua/harkivska-oblast-2>. Дата звернення: Вер. 15, 2022.
- [5] В. О. Денисенко, О. Ю. Мельников, «Постановка задачі розробки програмного забезпечення для визначення незаконної вирубки лісу», *Молодь і наука: виклики та перспективи: збірник тез наукової конференції молодих вчених 16 грудня 2021р.*, Краматорськ: Донецька обласна державна адміністрація, Рада молодих вчених при Донецькій облдержадміністрації, 2021, с. 134-135
- [6] О. Ю. Мельников, В. О. Денисенко, «Розробка програмного забезпечення для виявлення незапланованої вирубки лісу», *Автоматизація та комп'ютерно-інтегровані технології у виробництві та освіті: стан, досягнення, перспективи розвитку: матеріали Всеукраїнської науково-практичної Internet-конференції*, Черкаси, 2022, с. 55-57
- [7] Співаківка (Ізюмський район). Вікіпедія. [Електронний ресурс]. Режим доступу: [https://uk.wikipedia.org/wiki/Співаківка_\(Ізюмський_район\)](https://uk.wikipedia.org/wiki/Співаківка_(Ізюмський_район)). Дата звернення: Вер. 15, 2022.

Application of Virtual Simulation for Device Design

Viktor Bilishchuk

*Department of Metrology and Information-Measuring Technology
Ivano-Frankivsk National Technical University of Oil and Gas
Ivano-Frankivsk, Ukraine*

Abstract— The relevance of using virtual modeling of devices for engineering design is considered. An analysis of the publication on the application of computerized design technologies in the processes of device development was made. The possibilities and ways of using computer modeling in the process of developing devices in the Proteus Design and SolidWorks software environments are indicated. In the Proteus Design software environment, you can simulate the operation of electrical circuits containing programmable devices: microcontrollers and microprocessors, as well as design a printed circuit board. A large number of virtual tools are used to simulate the circuits, which simulate the operation of measuring devices, which facilitates the development of electrical circuits. The ARES PCB design module has an autorouter that allows you to automate the design process. As a result of the design of the printed circuit board, a pattern of tracks is obtained, which is used for the manufacture of the printed circuit board by various methods. On the basis of the comparison of the simulation results and the operation of the manufactured device, conclusions are made about the effectiveness of the application of simulation in the engineering design of electronic measuring devices.

Keywords— *Virtual Device, Three-Dimensional Model, Electrical Circuit, Automated Design System.*

I. INTRODUCTION

The development of new products is an important tool for the growth of companies. Engineering design is a science-intensive activity. The application of virtual simulation with the help of computer software greatly facilitates engineering design and has become a necessity in the present time. Virtual modeling saves time for the development of new products, reduces the number of physical prototypes and experiments, reduces the cost of production, and also increases the quality of final products, facilitates the understanding of the design of a new product, and improves management and automation processes.

II. RELATED WORK

Computerized design technologies have been used in the development of devices for a long time. A large number of scientific and scientific-methodological works and studies are devoted to the study of the problem of using information technologies for the design of devices. These works consider various aspects of the problem. The works [1-3] consider the concept of virtual and digital modeling, discuss the historical development and technologies, explore the consequences and advantages of this technology, describe in detail the specific branch applications of virtual modeling, and explore the advantages of analytical and virtual modeling. Approaches to the design of products and processes are investigated in [4]. These studies show how product data is used to analyze a variety of manufacturing activities in a virtual environment: product development, production planning, assembly analysis, work research, workplace design, work simulation, and production layout. It describes the methodology adopted and discusses the advantages and disadvantages of using the software.

The purpose of the study is to analyze the possibilities and ways of using computer modeling in the process of developing electronic measuring devices using the example of Proteus Design and SolidWorks software environments.

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

III. METHODOLOGY

Specialized tool environments are software tools designed for modeling, including educational construction from ready-made basic models. This includes software packages for the automated design of electronic circuits Proteus, NI Multisim, Micro-Cap, etc. These software tools allow you to create electrical circuits using graphic editors, edit component parameters, and connect "virtual" devices (generators, various meters, etc.). Analog, digital, and analog-to-digital devices can be simulated in these environments. These software products of computer modeling use "virtual" devices, by which we mean a mathematical model that imitates the structural and functional principles of the device.

Consider the Proteus Design package, which is a circuit modeling system based on models of electronic components adopted in PSpice. A distinctive feature of the Proteus Design package is the ability to simulate the operation of programmable devices: microcontrollers, microprocessor systems, DSP, etc. The Proteus Design package consists of two modules: ISIS - a graphic editor of electronic circuits with the function of modeling created circuits and ARES - a graphic editor of printed circuit boards with an autorouter, which makes Proteus Design an automated design system.

In ISIS, you can simulate the operation of electrical circuits in two ways: in real time and with the construction of graphs. The possibility of performing simulation in real time strongly depends on the complexity of the scheme and the computing capabilities of the computer. By modeling with the construction of graphs, it is possible to perform a more detailed analysis of circuits, which cannot be carried out in real time (construction of transmission, amplitude-frequency, voltage-amperage characteristics).

When performing interactive simulation of circuits, there is an opportunity to interact with the virtual circuit as with a real device. At the same time, virtual tools, devices, components and indicators are used. Virtual devices include: voltmeter, ammeter, oscilloscope, counter/frequency meter, signal generator, logic analyzer, etc. Virtual components are used to influence the scheme: switches, potentiometers, thermal and photoresistors, thermocouples, etc. Interaction with such components is possible by clicking with the mouse on the image of the component itself. Both in the real device and in the virtual one, there are also various models of indicators and boards: LEDs, seven-segment and liquid crystal symbol boards.

A feature of the ISIS module is interactive simulation of device circuits with microprocessors. To simulate the operation of a microprocessor, a binary program file is loaded into the microprocessor model. In the future, you can debug the microprocessor program step by step. In the step-by-step program execution mode, various windows with information about the current state of the microprocessor model can be called up.

An electronic unit with a printed circuit board can be designed on the basis of the assembled electrical circuit of the device. Designing takes place by means of modeling the ARES module in two-dimensional space with the subsequent generation of a three-dimensional model. Features of printed circuit board and block design environments are the use of specialized layers for placing 2-D graphic elements that reflect the elements of the electronic block - radio elements, printed conductors, contact pads, etc.

Modeling a printed circuit board has two main stages: placement of radio elements and tracing of printed conductors. The procedure for placing elements on the printed circuit board in the ARES module consists in specifying their location with the mouse pointer and using the commands "rotate", "shift", etc. After placing the elements, trace the printed conductors. In the ARES module, printed conductors can be traced manually or with the help of an autorouter - a special program. The most convenient is the initial tracing by an autorouter followed by manual adjustment of the configuration of the printed conductors.

The three-dimensional model of the printed block can be used for three-dimensional modeling of the device in the SolidWorks environment, which involves modeling the case, placing the parts of the device and the actual printed block in it. Three-dimensional modeling of the device consists in the composition of models of parts in the assembly - their placement and the creation of geometric relationships between them.

IV. RESULTS

The conducted research has the following results: the main stages of designing electronic devices are considered, the advantages of virtual modeling for the development of new electronic devices are highlighted.

V. DISCUSSION

The use of virtual modeling opens up many new possibilities in the design of new products, which is quite difficult, time-consuming or expensive with traditional approaches. The main advantages of using virtual technologies in engineering can be distinguished: correction of design errors before the creation of a layout, the ability to visually see the product even before it is created, significant time savings, less effort and resources spent on project development.

VI. FUTURE RESEARCH

This work does not outline all aspects of the research, therefore such problems as: research on the compatibility of different modeling environments to increase the efficiency of product design, application of finite element methods for the analysis of the design of product elements need to be studied in the future. These and other issues remain a promising topic of future research, and may form directions for further work taking into account the obtained results.

VII. CONCLUSION

The conducted analysis makes it possible to conclude that virtual modeling has prospects for increasing the efficiency of new product development processes, in particular measuring devices. The use of the Proteus Design and SolidWorks environments in the development of devices is promising both for the design of devices and for the study of subjects related to the study of the principles of device design by students of education.

REFERENCES

- [1] Availability and Capabilities of 'Low-End' Virtual Prototyping. Prime Faraday Technology Watch – 2001. [Online]. Available: https://www.lboro.ac.uk/microsites/mechman/research/ipm-ktn/pdf/Technology_review/virtual-prototyping-early-in-the-design-cycle.pdf. Accessed on: Feb 08, 2022.
- [2] Create 3D models with Virtual Reality. Ghost. 2019. [Online]. Available: <https://www.afternow.io/create-3d-models-with-virtual-reality/> Accessed on: Feb 08, 2022.
- [3] Bill Jennings and Kevin Bourne. Virtual Prototyping Vs. Physical Prototyping. Communication technologies Inc. 2019. [Online]. Available: <https://stonemountaindesign.com/images/pdfs/TCTMay01-Proto.pdf>. Accessed on: Feb 08, 2022.
- [4] D.S.K. Chan, "Simulation modelling in virtual manufacturing analysis for integrated product and process design," *Assembly Automation*, vol. 23, no. 1, pp. 69–74, 2003. doi: 10.1108/01445150310460114. [Online]. Available: <https://www.emerald.com/insight/content/doi/10.1108/01445150310460114/full/html>.

The web application of calculation of parameters for a small hydroelectric power plant

Iryna Vashchysyak¹, Serhii Vashchysyak²

¹*Department of Energy Management and Technical Diagnostics, Ivano-Frankivsk National Technical University of Oil and Gas*

²*Department of Information Technology, King Danylo University
Ivano-Frankivsk, Ukraine*

Abstract—In the mountain regions of Ukraine, there is a problem with the power supply for remote facilities, which are located in the gorges near the channels of small rivers. Gravity-rotating small hydropower plants, which can operate at minimum water pressures, are considered as the best suited for such conditions. The existing method of calculating parameters of gravitational-rotating small hydropower plants provides series of multiple procedures for recalculating parameters to obtain optimal values of its characteristics. To eliminate this drawback, a web application has been developed that optimizes and automates the calculation process. Optimization is carried out by selecting the dynamic characteristics of the vortex to organize the work of the impeller of a certain type of turbine with the maximum possible efficiency for it. The web application is written in JavaScript using the React library and the MongoDB database.

Keywords—*Gravitation Water Vortex Power Plant, Method Of Calculation, Generation Of Electricity, Small Hydroelectric Power Stations, Web Application*

I. INTRODUCTION

The aim of the work is to create a web application to optimize the calculation of the parameters of a small hydropower plant, the choice of the type of impeller of the turbine and the selection of the optimal model of the hydraulic generator from the database. The developed web application will allow to develop small hydropower projects for mountain rivers with insignificant hydro potential in a short time, which will promote electrification of objects remote from the infrastructure.

II. RELATED WORK

In the mountain regions of Ukraine, there are problems with electricity supply to remote facilities, especially those located in mountain gorges near small riverbeds. The problem can be solved by gravitational-rotating small hydropower plants, which operate on the principle of vortex formation [1]. Part of the river runoff is used for the operation of such small hydropower plants, and the hydraulic working pressure can be minimal. Gravitational-rotating small HPPs begin to generate electricity at a head of 0.7 m.

The method of calculating the parameters of gravitational-rotating HPPs is given in [2]. The essence of this technique is an iterative procedure for calculating the electrical and hydraulic parameters of a small hydropower plant with sequential change of some of them to achieve optimal characteristics of the turbine and generator at a given river hydropower potential.

III. METHODOLOGY

The paper [2] considers the methods of placing the hydroturbine in the vortex of the gravitational-rotating HPP and indicates that to ensure its minimum size and obtain high efficiency, the impeller of the hydroturbine should be placed in the lower part of the vortex. The process of vortex formation and the way to improve its dynamic characteristics are described. In accordance with this, a method of calculating the hydraulic parameters of the fluid jet, turbine and vortex, which are associated with the electrical power of the AC generator.

The paper [2] shows the formulas for calculating the parameters of the gravitational-rotating HPP. The water flow through the cross section of the bottom hole for the calculation is selected not more than 25% of the average annual runoff of the river, and the hydraulic pressure is selected from the height difference along the riverbed at the installation site of a small hydropower plant. The optimization of the calculation is to select the dynamic characteristics of the vortex to obtain the maximum efficiency of the turbine with an impeller of a specific size within the specified limits of pressure change. Optimization is also carried out in order to reduce the size of the impellers of hydraulic turbines while obtaining maximum power generation. In both cases, you need to constantly list most of the parameters of a small hydropower plant to achieve the desired result.

IV. RESULTS

To automate this process, better understand the impact of parameters on each other, automatic selection of the type of hydro turbine, selection of a small hydroelectric power plant model and ease of use, a web application was created in the JavaScript programming language [3]. The web application uses the React library and the MongoDB database.

The database of the web application includes the main types of impellers of hydraulic turbines (with visual images) and their efficiency, some constants, as well as models of ready-made small hydropower plants that are offered on the market.

In general, the web application consists of 12 blocks of formulas and a navigation block common to them all, a block of turbine model selection and a block of calculation results for all parameters. In the navigation block there are symbols of the parameters that are calculated.

In the first block of formulas at the top is the name of the web application "Online power calculator of hydropower plants" (Figure 1). Below it is the name of the parameter that is calculated "Coefficient of compression of the liquid jet". The calculation formula and the variables included in it are given below. The web application provides two ways to enter variables: manually and with the arrows located on the right side of the field to enter a variable. Arrows appear when you hover over them. In the left part of the formula blocks there is a navigation block, selecting the variable from which the automatic transition to the block with the desired formula is performed. Below the input fields of the variables is the inscription "Result", under which the calculated value of the variable appears after clicking on the "Calculate" button.

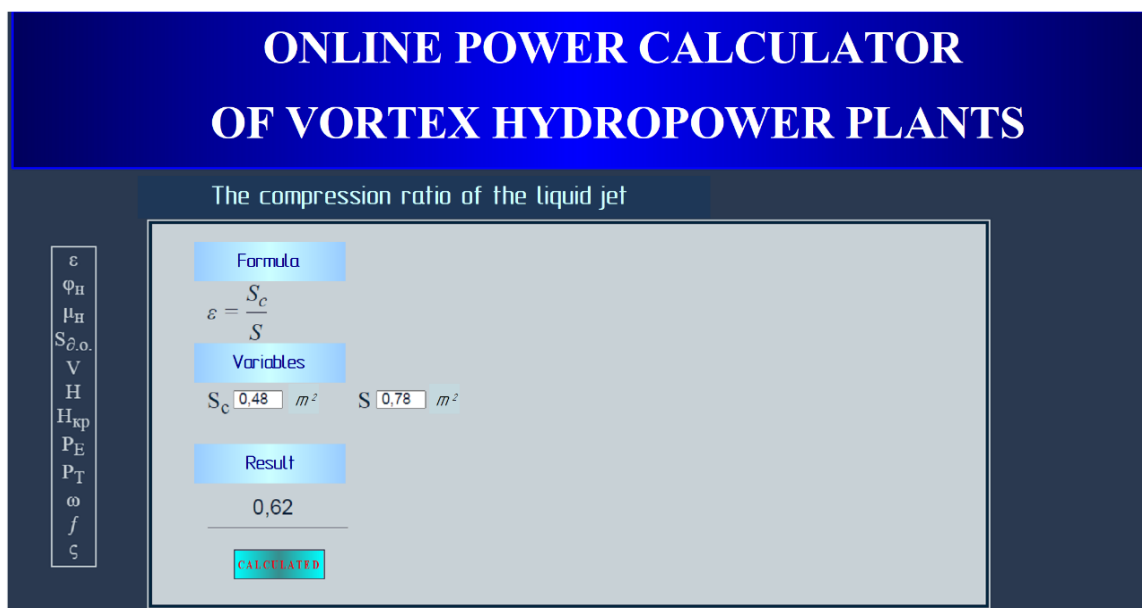


Figure 1. View of the first block of formulas

All formula blocks, starting with the second, have fields with a check box (Checkbox), in which you can use either the recommended values to optimize the parameters of the vortex, or entered manually.

In the lower right part of the block of formulas "Power developed by the hydroturbine" there is a scale, moving the slider of which it is possible to change efficiency of the hydroturbine. To the right of the scale appears an image of a hydroturbine for a small hydropower plant (Francis, Kaplan, Thomson, Turgot, Pelton), which best corresponds to the selected efficiency (Figure 2).

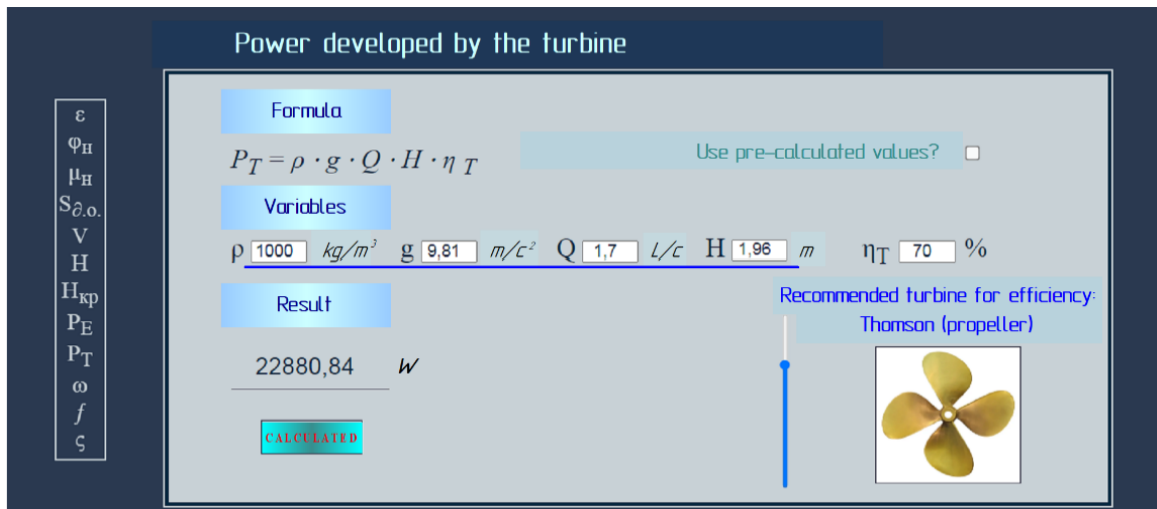


Figure 2. View of the block of formulas with the choice of the type of turbine

The turbine model selection unit "Characteristics of the recommended turbine" optimizes the parameters of a small hydropower plant according to the following calculation results: water flow (Q), turbine power (Pt), diameter (D), generator power (Re), weight. When you click on any of the buttons on the right, a parameter scale will appear, from which the closest to the calculated value is selected. After that, the program recalculates the data to optimize the dynamic characteristics of the vortex for the selected efficiency of the turbine and offers a model of a small hydropower plant, which is best suited for the calculated parameters. The table "Characteristics of the recommended turbine" provides the main technical data of the proposed model (Figure 3).

Characteristics of the recommended turbine		Generator power (kW)			
Turbine model	ZD760-LM-60	<ul style="list-style-type: none"> • 9 • 12 • 15 • 25 • 34 • 46 • 8 • 10 • 40 • 50 • 55 • 60 • 75 • 80 • 100 			
Water flow	1,2 ± 0,1				
Turbine power	30 kW				
Reduction of a reducer	variable (5-25)				
Hole diameter	60 sm				
Generator model	PM-2.0				
Electric power	25 kW				
Phase	3-phase				
Power factor	0.8				
Protection	IP23				
Temperature	30-50				
Humidity	90				
Gross weight	1510 kg				
Choose a turbine for:					
Water flow (Q)	Turbine power (Pt)	Diameter (D)	Generator power (Pe)	Weight	DOWNLOAD RESULTS

Figure 3. View of the turbine model selection unit

The table "Calculations" of the block of calculation results (Figure 4) displays the values that the user enters and calculates, which are automatically updated when you click "Update". After clicking the "Download calculation results" button, all calculations and data of the selected small HPP model will be saved on the computer in text format.

CALCULATIONS	
Parameters	Value
ε	0.62
φ_H	1.04
μ_H	0.64
$S_{\partial.o.}$	0.48 m ²
V	3.56 m/c
H	1.96 m
$H_{кр}$	0.45 m
P_E	21597.86 kW
P_T	22880.84 kW
ω	25.25 rad/c
f	4.02 Hz
ζ	12.50
η_T	70 %
Recommended turbine	Thomson (propeller)

[DOWNLOAD CALCULATION RESULTS](#)
[Update](#)

Figure 4. View of the block of calculation results

V. DISCUSSION

Using the developed web application will allow you to automate the design process, carry it out in a short time with maximum efficiency and save money on equipment. This will speed up the electrification of remote mountain facilities.

VI. FUTURE RESEARCH

The developed web application is easy to use, uses a small amount of computer memory and can be easily supplemented with new settings and data.

VII. CONCLUSION

A web application has been created that optimizes the calculation of the parameters of a small gravitational-rotating HPP for operation in a vortex with dynamic characteristics that provide maximum efficiency of the turbine. The web application has a user-friendly interface and offers a choice of ready-made models of small hydropower plants that are sold on the market and are the closest in parameters to the calculated ones. Selection of models of small hydropower plants is possible by water consumption, impeller diameter and power of the turbine and generator.

REFERENCES

- [1] “Нова водоворотна турбіна зробила гідроенергетику “зеленою”, *ЕкоТехніка*, 2018. [Online]. Available: <https://ecotechnica.com.ua/energy/voda/3137-novaya-vodovorot-naya-turbina-sdelaet-gidroenergetiku-potuzhnosti-vyrovykh-hes>.
- [2] I. Vashchyshak, “Designing gravity-rotary hydroelectric power plants for small rivers of the Carpathian region”, *Scientific Bulletin of UNFU*, vol. 31(2), pp. 93-97., 2021. doi: 10.36930/40310215.
- [3] Onlain kalkuliator potuzhnosti vyrovnykh HES, 2020. [Online]. Available: <https://infinite-brook-92538.herokuapp.com/#ElectricPowerGenerator>.

Modeling Technologies for Assessment of Information Security Risks in e-Commerce

Olena Hladka, Ivan Karpovich, Daria Buryan

*Department of Computer Technology and Economic Cybernetics
National University of Water and Environmental Engineering
Rivne, Ukraine*

Abstract—The paper is devoted to the study of cybersecurity of information systems at an enterprise in the e-commerce. We performed modeling, analysis and assessment of information security risks. A methodology for risk assessment is proposed, which takes into account the criticality of business processes, the differentiation of the significance of assets, as well as the impact of the moment of realization of possible threats on the consequences for the business process and related damage. Recommendations for improving the efficiency of the information protection system are formulated.

Keywords—*Cybersecurity, E-commerce, Business Process, Criticality, Threat, Information Protection, Information Security Risk Analysis, Risk Assessment Model.*

I. INTRODUCTION

The dependence on information in the business environment is extremely high, because many trade transactions are carried out electronically over the Internet. Such information dependence has led to a significant increase in the impact of the security level of information systems on success, and sometimes simply on the possibility of doing business.

The main criterias for evaluating the information security of e-commerce systems are:

- *availability* is resistance to failure, resource utilization, updates after failures, variability;
- *integrity* is impossibility of changes, blocking or destruction;
- *confidentiality* is trust, formality, preservation, lack of hidden access;

as well as additional ones are *manageability, relevance, guarantee of performance* (delivery), *importance*, etc.

The psychological factor associated with awareness of the threat of potential fraud remains the main obstacle to using the Internet as a means of conducting commercial transactions. Users and professionals still do not consider the Internet as a safe environment. Surveys show that the greatest potential threat is the unauthorized acquisition of personal data when using open Internet communication channels. According to the developers of the VISA payment system, about 23% of e-commerce transactions are not carried out due to the client's fear of entering personal information when working, for example, with an electronic store [1].

II. RELATED WORK

In order to avert a threat to information security, according to literary sources [1-3], it is necessary to fulfill the following conditions:

- exclude the possibility of interception of personal or bank information during the transaction;
- exclude the possibility of extracting this information from databases;
- eliminate the possibility of using "stolen" information for your own purposes.

There are many e-commerce protections that help create a secure environment for e-commerce operations. New approaches are being developed that involve the creation of a large number of complex encryption algorithms, special programs for intercepting and neutralizing attacks. However, it is currently impossible to achieve complete security of e-commerce systems, which still contain many vulnerabilities [4].

Information Security (IS), as protection against harm or a system of measures to reduce it, involves the preservation of the system's stability, invulnerability and the possibility of self-development. In e-commerce, information security is one of the most important components of doing business. The main object of IS is the company's commercial information. Therefore, before organizing business activities on the Internet, it is necessary to take care of a well-thought-out information protection system.

III. METHODOLOGY

The fundamental basis for building any system of protection against cyber threats is risk analysis. Risk analysis is the procedure for identifying risk factors and assessing their significance. Thus, risk is the probability that certain undesirable events will occur that will negatively affect the achievement of the goals of a particular business process. Risk analysis includes both risk assessment and methods of reducing them or reducing the associated adverse consequences. In information security, risk assessment techniques have appeared with the aim of forecasting the possible damage associated with the implementation of threats and estimating the required amount of investment for the construction of information protection systems.

Risk analysis can be divided into two complementary types: qualitative and quantitative. Qualitative risk analysis aims to determine (identify) factors, areas and types of risks. Quantitative risk analysis should make it possible to numerically determine the size of individual risks and the overall size of the risk as a whole. The results of qualitative risk analysis, in turn, serve as basic information for conducting quantitative analysis. However, quantitative assessment is complicated by the fact that calculations require reliable input information and a clearly defined assessment scale.

The IS risk assessment process is carried out as follows. At the first stage of risk identification, information assets are determined and the value of assets is established. Based on the analysis and determination of the probability of threats, the characteristics of the classes of threats that lead to the realization of the risk are formed. Identification of vulnerabilities of information assets, determination of the degree of vulnerabilities is carried out.

At the second stage, by calculating the probability of the occurrence of an event related to the implementation of threats (exploitation of vulnerabilities), the identified threats are evaluated, that is, the probability that the threat will be implemented is determined. By combining the importance of information assets and the possibility of incidents, the value of the information security risk for the information asset is calculated.

The probability of threat realization can be estimated using the formula [5]:

$$P = (1 - P_p) \cdot (1 - P_D) \cdot (1 - C),$$

where P (Probability) is the probability that the threat will be realized within a certain period, for example, a year; P_p (Probability of protection) is the probability of preventing a threat; P_D (Probability of detection) is the probability of detecting a threat; C (Complexity) is the difficulty of implementing the threat (the probability that the threat will not be implemented due to the difficulty of implementation).

The predicted frequency of threat realization (F) is determined by the formula:

$$F = P \cdot F_o,$$

where F_o (Frequency of occurrence) is the predicted frequency of occurrence of the threat per year.

Then, the predicted level of losses for the year can be calculated using the formula:

$$L = D \cdot F,$$

where L (Losses) is the level of losses from the implementation of threats; D (Damage) is the expected value of the loss from the threat for the year; F (Frequency) is the predicted frequency of threat implementation per year.

Determining a reasonable level of expenditure of various resources to create the security of the functioning of the considered systems in the conditions of their inherent nonlinearity, poor predictability of behavior and traditional resource limitations is an important and ambiguous task. In modern studies, the concept of a critically important object is widely used, for example, critically important information objects, critically important objects of information infrastructure, etc [6].

Let's consider a number of important risk-oriented features of e-commerce business processes that are used to assess IS risks. First, it is the criticality (importance) of the business process, which is determined by the extent of its contribution to the achievement of the company's strategic goals. A business process whose criticality is high and the assets associated with it should have a higher priority during the organization of security measures than business processes with a lower criticality (other things being equal). Then we take into account the value of the asset in relation to a specific business process. Finally, the implementation of all stages of each business process requires some time, and the business process itself can be iterative in nature. The stage (iteration) time, predicted or calculated based on the real situation, is the normative time for passing all stages of the business process from the first action to obtaining the result, that is, achieving the goal of the business process.

The K_1 criticality factor of each business process is defined as the share of the number of critical success factors related to this business process in the total number of critical success factors.

The accounting of the time factor is carried out according to two components: the coefficient of criticality of deadlines (for the business process) and the coefficient of impact on the deadline (for the IS threat) [7]. The criticality coefficient of terms is determined for each business process on a quality scale. The main evaluation criterion is financial losses or other equivalent consequences for the firm (reputational risks, human resource costs) caused by a change in the time of iteration of the business process. The coefficient of impact on the term is determined for each threat. Note that each threat affects the business process differently, including the time factor, in our case, the change in the time of the stage of the business process.

For each threat, the K_2 coefficient of influence on the execution time can be presented in the form [7]:

$$K_2 = (T + \Delta T)/T,$$

where T is the execution time of one stage (iteration) of the business process; ΔT is the deviation in the execution time of one stage of the business process.

The K_3 criticality coefficient of terms is determined as follows:

$$K_3 = K_1 \cdot K_2.$$

The found K_3 coefficient of criticality of the terms allows to submit the d coefficient of damage (destruction), which expresses the extent of the destructive effect of the threat on the asset or assets, in the form of a function $d(K_3)$ and to calculate the probable losses from the implementation of a separate threat using the formula for calculating the cost of risk [8]:

$$R = \sum_{k=1}^{k_1} F \cdot P \cdot d(K_3) \cdot c(a_k).$$

Here R is the cost of risk; k_1 is the number of assets targeted by the threat; F is the frequency of threat occurrence; P is the probability of the occurrence of a threat; $d(K_3)$ is the coefficient of damage (destruction); $c(a_k)$ is the value of the asset a_k , $a_k \in A$, where A is the set of assets or resources targeted by the threat.

The list of assets for each business process is carried out separately at the stage of assessing the value of assets (information, resources, etc.). Note that the value of the same asset in the context of different business processes may be different.

IV. RESULTS

In order to reduce the likelihood of threats being realized, it is necessary to take measures that increase the likelihood of preventing the threat. In particular, protection of information from interception is carried out using SSL (Secure Sockets Layer) and SET (Secure Electronic Transaction) encryption protocols. In addition, participants of Internet commerce use such methods of identification of the card holder as verification of CVV2 / CVK2 codes, etc.

With regard to risk management for all major information systems of electronic business, an analysis and assessment of existing threats related to violations of the availability / integrity / confidentiality of information systems, which can lead to significant damage and impact on business processes, should be carried out regularly. In order to level the consequences of dangerous threats and ensure the timely restoration of the most important business operations, plans for the restoration of the relevant technological sections of the information infrastructure must be developed. The construction of an information security risk management system as a component of the information security management system aims to minimize losses that may be caused by violations in information systems [9]. Based on the analysis of risks and threats, planning, implementation and control of countermeasures aimed at protecting IS is carried out, in particular, using modern methods of digital data exchange in conditions of strong interference [10].

V. CONCLUSION

The developed method of analysis and assessment of information security risks in the field of e-commerce allows you to differentiate the importance of assets, assess the criticality of business processes and calculate the cost of risk, that is, the potential damage that the implementation of possible threats can cause. The presented recommendations for increasing the efficiency of the information protection system involve the implementation of both technical (hardware and software) and organizational measures.

REFERENCES

- [1] S. G. Arbuzov, Yu. V. Kolobov, V. I. Mishchenko, and S. V. Naumenkova, *Visa // Banking Encyclopedia*, Kiev: Center for Sci. Research of the National Bank of Ukraine: Knowledge (in Ukrainian), 2011.
- [2] V. V. Vitlinsky, and G. I. Velikoivanenko, *Risicology in Economics and Entrepreneurship: Monograph*, Kiev, Ukraina: KNEU (in Ukrainian), 2004.
- [3] O. Altukhova, R. Matulevičius, and N. Ahmed, "An Extension of Business Process Model and Notation for Security Risk Management", *International Journal of Information System Modeling and Design*, No. 4(4), pp. 93-113, 2013. doi: 10.4018/ijismd.2013100105
- [4] V. Mokhor, O. Bakalynskiy, and V. Tsurkan, "Analysis of information security risk assessment representation methods", in *Collection "Information Technology and Security"* (in Ukrainian), No. 6(1), pp. 75-84, 2018. doi: 10.20535/2411-1031.2018.6.1.153189
- [5] I. V. Ruban, V. V. Tyutyunyk, V. I. Zabolotny, and O. O. Tyutyunyk, "Assessment of the vulnerability of cybersecurity objects based on a risk-oriented approach", *Information Security* (in Ukrainian), Vol. 26, № 3, pp. 145-155, 2020. doi: 10.18372/2225-5036.26.14947
- [6] ISO/IEC 27032:2012(E), *Information technology — Security techniques — Guidelines for cybersecurity*, 2012.
- [7] Martin Fowler, *UML Distilled: A Brief Guide to the Standard Object Modeling Language*, 3rd Edition, Addison-Wesley Professional, 2005.
- [8] I. Karpovich, O. Hladka, and Y. Bukhalo, "Technologies of modeling and assessments of the information security risks", *Technical sciences and technologies: science journal*, Chernihiv, Ukraina: Chernihiv Polytechnic National University (in Ukrainian), No. 1 (23), pp. 62-68, 2021. doi: 10.25140/2411-5363-2021-1(23)-62-68
- [9] Scott Donaldson, Stan Siegel, Chris Williams, and Abdul Aslam, *Enterprise Cybersecurity: How to Build a Successful Cyberdefense Program Against Advanced Threats*, Apress, 2015.
- [10] Lazarovych, I., Kozlenko, M., Kuz, M., ... Savka, I., Pikuliak, M. "Software Implemented Enhanced Efficiency BPSK Demodulator Based on Perceptron Model with Randomization". *IEEE 3rd Ukraine Conference on Electrical and Computer Engineering, UKRCON. Proceedings*, 2021, pp. 221–225. doi: 10.1109/UKRCON53503.2021.9575458

Роль навігаційних систем в організації ефективної роботи автомобільного транспорту

Михайло Музикін¹, Галина Нестеренко¹, Олег Стрелко²,
Ніна Клочкова³

¹Український державний університет науки і технологій
Дніпро, Україна

²Державний університет інфраструктури і технологій
Київ, Україна

³Університет митної справи та фінансів
Дніпро, Україна

Анотація—Одним з основних способів підвищення ефективності роботи автомобільного транспорту України є застосування навігаційних систем та їх розвиток. Метою роботи є дослідження ролі навігаційних систем в організації ефективної роботи автомобільного транспорту України.

Ключові слова—Автомобільний транспорт, Навігаційна система водія, Супутниковий моніторинг.

I. ВСТУП

Ефективність роботи автомобільного транспорту – важливий чинник, що впливає на розвиток економіки країни, відіграє важливу роль під час воєнних дій, сприяє підвищенню культури та освіченості населення, підвищує ефективність праці, а також має велике політичне значення. Одним з основних способів підвищення ефективності роботи автомобільного транспорту України є застосування навігаційних систем та їх розвиток. Мета роботи – дослідити роль навігаційних систем в організації ефективної роботи автомобільного транспорту України.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Дослідженням питання розвитку навігаційних систем присвячено цілий ряд наукових праць [1-3]. Проведений аналіз публікацій з цієї тематики доводить актуальність нашого дослідження.

III. ПРИЗНАЧЕННЯ ТА ФУНКЦІОНАЛ НАВІГАЦІЙНИХ СИСТЕМ

Навігаційні системи пропонують широкі можливості для організації ефективної роботи транспортної компанії або іншої організації, що використовує автомобільний транспорт. Широке розповсюдження отримали системи супутникового моніторингу транспорту, які вирішують проблеми та завдання, пов'язані з підвищенням ефективності роботи автотранспорту. Супутниковий моніторинг – система на основі технологій супутникової навігації, що дозволяє відстежувати основні параметри транспортного засобу на маршруті, вирішувати проблеми з нецільовим використанням автомобілів, оптимізувати їх роботу, знижуючи витрати і підвищуючи ефективність бізнесу в цілому. Система реалізується у вигляді бортового обладнання (трекера, радіопередавача, системи датчиків), яке проводить вимірювання найважливіших параметрів, і по радіоканалу (зазвичай використовуються мережі мобільного зв'язку або окремий радіоканал) передаються на сервер і термінал диспетчера (оператора). Ця інформація відображається на комп'ютері диспетчера (а також

записується в GPS або ГЛОНАСС-трекер на автомобілі), що і забезпечує постійний моніторинг.

Основні функції системи супутникового моніторингу:

1. Відстеження поточних координат транспортного засобу. Це дає можливість відразу виявляти відхилення від маршрутів, несанкціоновані зупинки і нецільове використання автомобіля.

2. Відстеження витрати палива і порівняння фактичних даних з паливними картами (завдання вирішує спеціальний датчик витрати палива, що встановлюється в баку). Це дозволяє будувати максимально ефективні маршрути, а також запобігати зливу палива.

3. Відстеження основних параметрів вантажу та автомобіля в цілому.

4. Вирішувати проблеми при виникненні надзвичайної ситуації (в разі нападу на автомобіль або при викраденні).

На сьогодні в Україні забезпечена можливість використання глобальної системи супутникового моніторингу: GPS (США). Регіональні системи Галілео (Європейський союз), Бейдоу (Китай) можуть бути доступними у перспективі.

NAVSTAR GPS — високоточна супутникова система навігації, яка дозволяє визначити місцезнаходження об'єкта, його широту, довготу та висоту над рівнем моря, а також напрямок і швидкість його руху. Комплекс NAVSTAR розроблений, утілений і належить Міністерству оборони США. Використання GPS забезпечує високі надійність функціонування та точність даних, а також дає змогу інтегрувати в мережу велику кількість GPS-пристроїв: стільникових телефонів, навігаторів, GPS-трекерів, GPS-жучків тощо.

Галілео – система супутникової навігації Європейського Союзу та Європейського космічного агентства. Використання цієї НС забезпечує гарантований сигнал і систему попередження в разі зниження точності позиціонування та гарантує підвищену точність визначення місцеположення в межах 4 м в горизонтальній площині і 8 м у вертикальній.

З 27 грудня 2012 року в комерційну експлуатацію запущена регіональна система позиціонування BeiDou («Сузір'я "Велика Ведмедиця"»), яка охоплює територію Китаю і суміжних держав. Китай планує зі збільшенням кількості супутників довести її рівень до глобального під назвою Compass. Стверджується, що в даний час точність позиціонування в цій системі досягає 10 м. Навігатор в китайській системі є не тільки приймачем, але і передавачем сигналу. Станція моніторингу через два супутника посилає сигнал користувачеві. Пристрій користувача після отримання сигналу посилає відповідний сигнал через обидва супутника. Наземна станція по затримці сигналу розраховує географічні координати користувача, визначає висоту по наявній базі даних і передає сигнали на пристрій користувача сегмента.

Навігаційні системи діляться на дві групи: навігаційні системи водія і диспетчерські навігаційні системи.

Навігаційні системи водія (НСВ) призначені для вказівки за допомогою дисплея на приладовій панелі поточного місцезнаходження ПС водієві, прокладання найкоротшою траси маршруту, контролю встановленого графіка руху. Практично всі сучасні НСВ використовують для визначення місцезнаходження ПС систему GPS або спільно сигнал GPS і ГЛОНАСС, що дозволяє збільшити кількість одночасно видимих супутників і, відповідно, підвищити точність позиціонування. Найбільша точність НСВ досягається при її поєднанні з трасувальником. У цьому випадку неминучі похибки коригуються за умовою мінімуму середньоквадратичної помилки. Такі системи отримали назву інтегрованих систем GPS - Dead Reckoning GPS (DRGPS).

Диспетчерські навігаційні системи (ДНС) призначені для передачі даних про місцезнаходження ПС на диспетчерський пункт (в окремому випадку АТО).

IV. РЕЗУЛЬТАТИ

Використання системи диспетчерського управління сприяє:

- підвищенню ефективності роботи транспорту;

- зменшенню холостого пробігу автомобілів, так як місцезнаходження транспортного засобу відомо і з водієм може бути встановлений контакт в будь-який час;
- зростанню ефективності завантаження і часу доставки вантажів замовнику, так як диспетчер може оперативно реагувати на інформацію про наявність вантажу і вільного місця в транспортному засобі;
- зростанню ефективності роботи водія, так як він не повинен втрачати час на вимушене очікування завдань на перевезення, відхилятися від маршруту і шукати телефон для зв'язку (водій може стояти в черзі на митному пункті і одночасно тримати зв'язок зі своїм офісом замість того, щоб втрачати час на вимушену парковку і пошук телефону для зв'язку з диспетчером);
- зменшенню витрат на телефонний зв'язок в результаті більш ефективного виду зв'язку - передачі даних та економії часу при зайнятості телефонних ліній зв'язку або відсутності абонента на робочому місці.
- підтвердження про доставку вантажу одержувачу дається негайно, що дозволяє без зволікання виставити рахунок замовнику, а не чекати, коли транспортний засіб повернеться на базу. Крім того, замовнику може бути представлено доказ того, що вантаж перевозився у відповідності до вимог замовника;
- диспетчер може здійснювати безпосередній контроль стану вантажу і оперативно передавати цю інформацію замовнику;
- поліпшується планування ремонту і техобслуговування, так як керуючий транспортними засобами має більше інформації про стан транспортних засобів та розкладі їх руху в той час, коли машини перебувають ще далеко від бази;
- підвищується рівень обслуговування клієнтів шляхом їх інформування про процес транспортування і якісному стані вантажу під час його перевезення або прибуття автобусів на зупинний пункт. Це особливо важливо при транспортуванні цінних, швидкопсувних, небезпечних вантажів, для рефрижераторного парку та при доставці вантажів в точно встановлені терміни, при роботі автобусів в умовах щільного руху, на заміських трасах.

V. ОБГОВОРЕННЯ

Всі ці фактори дозволяють досягти головної мети будь-якої комерційно керованої діяльності – скоротити експлуатаційні витрати і підвищити дохід за рахунок оптимізації роботи транспорту і персоналу компанії.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальші дослідження доцільно проводити в напрямку адаптації використання технології доповненої реальності в навігаційних системах водія.

VII. ВИСНОВКИ

Використання навігаційних систем є необхідною складовою в організації ефективної роботи автомобільного транспорту, адже за їх рахунок забезпечуються:

- підвищення ефективності використання наявного парку транспорту і персоналу;
- зниження потреби в розширенні парку автотранспорту;
- зниження втрат від крадіжок вантажу і викрадень або нецільового використання автотранспорту завдяки удосконалюванню системи забезпечення безпеки;
- зменшення витрат на технічне обслуговування, паливо, мастильні матеріали за рахунок оптимізації маршрутів і зниження непродуктивного пробігу автотранспорту;
- поліпшення обслуговування клієнтів і можливість залучення нових клієнтів за рахунок розширення спектра послуг і оперативного реагування на запити;
- більш чітке перспективне планування роботи на основі об'єктивної інформації про реальний пробіг кожної одиниці автотранспорту і зниження втрат, пов'язаних з ремонтом і

простим автотранспорту, що особливо важливо для унікального автотранспорту і для машин з дорогим спеціальним устаткуванням;

– підвищення ефективності роботи персоналу і можливість уведення системи матеріального стимулювання, що базується на достовірній інформації про роботу кожного водія і заохочує до більш ефективного використання робочого часу, транспорту, паливо-мастильних матеріалів і спеціального устаткування.

ЛІТЕРАТУРА

- [1] R. Neumeier and G. Ostermayer, "Analysis of compass sensor accuracy on several mobile devices in an industrial environment," *International Conference on Computer Aided Systems Theory*, 2013, pp. 381-389.
- [2] D. Rettenmund, "Accurate visual localization in outdoor and indoor environment exploiting 3D image spaces as spatial reference," *International Archives of the Photogrammetry, Remote Sensing; Spatial Information Sciences*, vol. 1, 2018, ch. 42.
- [3] М. І. Музикін та Д. М. Солов'янов, "Дослідження дорожнього руху на території Республіки Казахстан," *Всеукраїнська науково-технічна конференція студентів і молодих вчених "Молода Академія 2022"*, 2022, с. 45-46.

Прогнозування якості програмного забезпечення

Борис Незамай

*Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація - Запропоновано новий метод прогнозування характеристик якості програмного забезпечення. Метод складається з наступних етапів: оцінка характеристик якості аналогічних попередніх проектів, встановлення впливу характеристик вимог на прогнозоване значення характеристик програмного забезпечення, верифікація вимог та кількісна оцінка прогнозних значень характеристик. Запропоновані етапи реалізуються методом експертного оцінювання.

Ключові слова – *Програмні засоби, Якість програмного забезпечення, Прогнозування якості, Верифікація вимог.*

I. ВСТУП

В даний час оцінка якості програмного забезпечення базується на цілому ряді міжнародних стандартів, та національних стандартів, які на них базуються, в тому числі і стандартів України. В її основі лежить метричний підхід на основі методів експертного оцінювання. Це відкриває певні можливості для спроб оцінити досяжні характеристики якості програмних засобів на етапі проектування та розроблення – тобто до прогнозування якості програмних засобів.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

В процесі огляду та аналізу існуючих опублікованих досліджень, які напряду чи опосередковано стосуються прогнозу якості програмного забезпечення, виділено наступні їх групи: аналіз ризиків, прогнозування помилок, та, власне, прогнозування характеристик якості програмного забезпечення. Описані спроби прогнозування характеристик якості програмного забезпечення містять один з трьох основних принципів: класичні методи прогнозування [1, 2], генетичні алгоритми [3, 4] або реалізації методів машинного навчання [5, 6, 7]. Останні дозволяють прогнозувати на основі встановлення різного роду класифікацій та співвідношення розроблюваного програмного засобу з одержаною класифікацією. Найпоширеніший інструмент при цьому – методи дерев рішень. В результаті аналізу виявлено наступні недоліки: описані підходи орієнтуються на вузькі аспекти оцінки якості, наприклад, лише на прогноз кількості помилок; методи прогнозування стосуються конкретних класів задач і мов програмування; методи прогнозування складні та дорогі в впровадженні і використанні, що не гарантує якість прогнозу.

III. МЕТОДИ

В процесі виконання роботи використано методи аналізу та синтезу. Для вирішення задачі прогнозування характеристик якості програмного забезпечення запропоновано використовувати методи експертних оцінок та формалізації.

IV. РЕЗУЛЬТАТИ

Встановлено, що одним з визначальних впливів на характеристики якості програмного забезпечення є якість вимог до його розробки [8, 9]. Запропоновано методику оцінки характеристик якості вимог на етапі їх верифікації за шкалою -1 (недостатньо), 0 (задовільно), +1 (повно) та методику визначення їх впливу на очікувані (прогнози) характеристики якості програмного забезпечення [10]. Дана методика передбачає встановлення зв'язків між характеристиками вимог (необхідність, повнота, узгодженість, однозначність, тестованість, відслідковуваність) та стандартизованими характеристиками якості програмного забезпечення методом експертних оцінок – приклад для нефункціональних вимог типу «Документованість» наведено на рисунку 1.

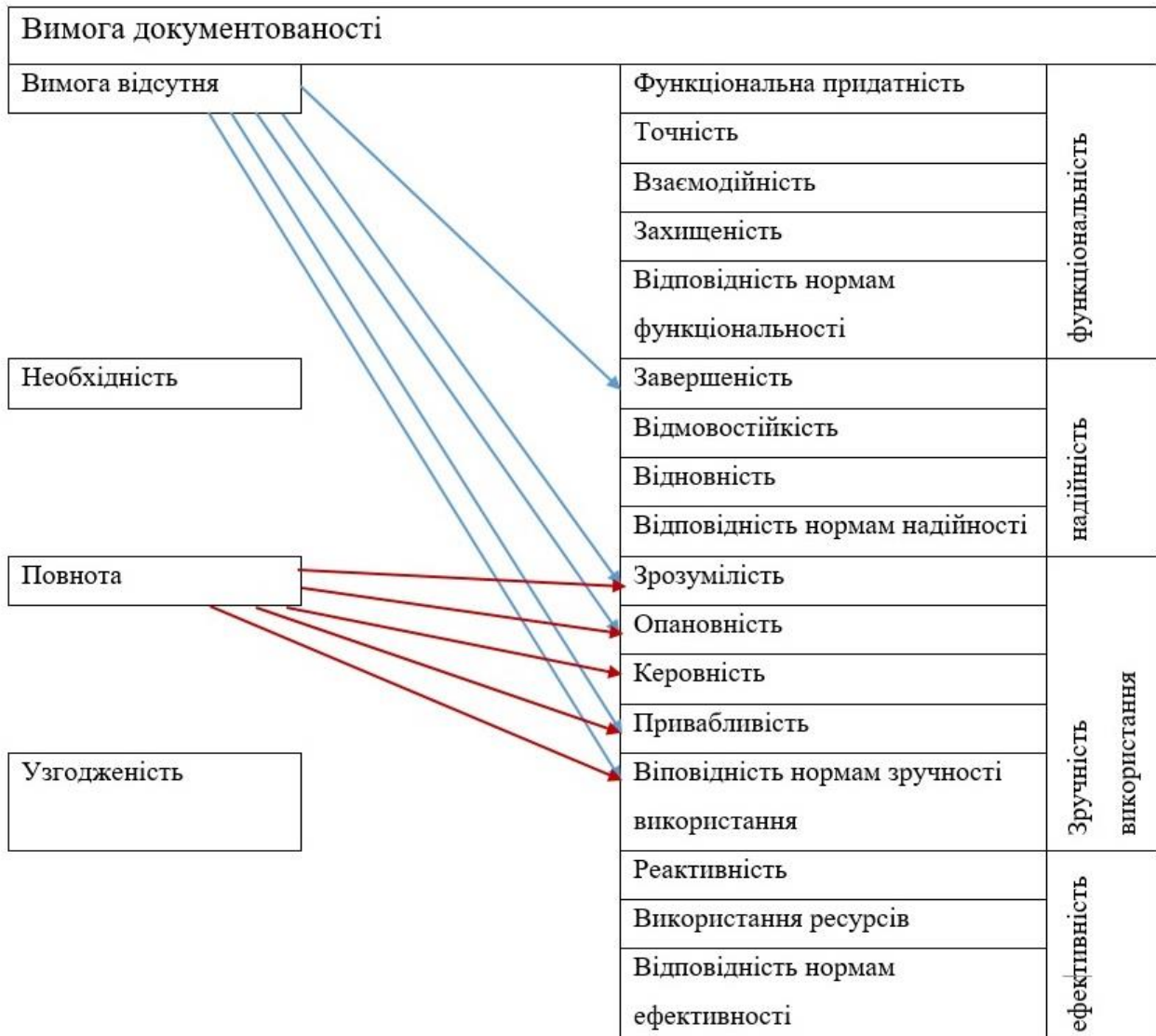


Рисунок 1 – Вплив характеристик вимог типу «Документованість» на прогноз характеристик якості

Результуючі прогнозні показники визначатимуться за формулою:

$$q = q_i + k_B \sum kpb + k_U \sum kpu + k_F \sum kpf + k_{NF} \sum kpnf , \quad (1)$$

де q – прогнозована метрика характеристики якості,

q_i – прогноз на основі історичного аналізу розробок команди, яка починає працювати над проектом, або прогноз на основі аналізу аналогічних продуктів,

k_B – ваговий коефіцієнт впливу на прогноз якості бізнес вимог,

k_{pb} – корекції прогнозу за бізнес вимогами,

k_U – ваговий коефіцієнт впливу на прогноз якості вимог користувача,

k_{pu} – корекції прогнозу за вимогами користувача,

k_F – ваговий коефіцієнт впливу на прогноз якості вимог до функцій,

k_{pf} – корекції прогнозу за вимогами до функцій,

k_{NF} – ваговий коефіцієнт впливу на прогноз якості нефункціональних вимог,

k_{pnf} – корекції прогнозу за нефункціональними вимогами.

При цьому пропонується шкалювання метрик та відповідних коефіцієнтів на $[0 \dots 1]$ з певним кроком, що встановлюватиметься експертами під конкретний проект і визначатиметься, в першу чергу, його об'ємом (типом і кількістю вимог).

V. ОБГОВОРЕННЯ

Очевидною користю запропонованого методу прогнозування є оцінка характеристик якості на основі наявного досвіду для подібних проектів та враховуючи наявні технічні вимоги на майбутній проект. При цьому, якщо значення характеристика знаходяться в околі нуля, можна превентивно встановлювати причини цього, що покращить загальні шанси на успіх проекту. Якщо значення прогнозу прямують до одиниці – можна зробити висновок, що високій якості майбутнього проекту ніщо не заважає. Головним недоліком запропонованого підходу є значна залежність результатів прогнозування від кваліфікації задіяних експертів.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В рамках запропонованого підходу до прогнозування характеристик якості програмного забезпечення перспективним є вивід відповідних залежностей для різних класів програмного забезпечення, розробка необхідного інструментарію, рекомендацій стосовно вибору коефіцієнтів, рекомендацій стосовно застосування даного підходу в індустрії.

VII. ВИСНОВКИ

В якості базової прийнято гіпотезу про визначальний вплив якості вимог на прогнозні значення характеристик якості програмного забезпечення. Запропоновано метричну оцінку характеристик вимог, запропоновано методику оцінки впливу характеристик вимог на очікувані характеристики якості та метод (формулу) для кількісної оцінки.

VIII. ПОДЯКА

Висловлюю щире подяку за наукове консультування та моральну підтримку науковому керівнику професору Миколі Кузю; та всій кафедрі Інформаційних технологій Прикарпатського національного університету імені В. Стефаника за надану можливість для професійного зростання.

ЛІТЕРАТУРА

- [1] Lanubile F. and Visaggio G. Evaluating predictive quality models derived from software measures: lessons learned. *Journal of Systems and Software*, 1997, Vol. 38, pp. 225–234.

- [2] Hong E., Wu C. Criticality models using SDL metrics set, *4 th Asia-Pacific Software Engineering and International Computer Science Conference*, 1997, pp. 23–30.
- [3] Evett M., Khoshgoftar T., Chien P. and Allen E. GP-based software quality prediction, *Third Annual Genetic Programming Conference*, 1998, pp. 60–65.
- [4] Zhang D., Tsai J.J.P. Machine learning and software engineering, *Software Quality Journal*, 2003, Vol. 11, Issue 2, pp. 87–119.
- [5] Ganesan K., Khoshgoftaar T., Allen E. Casbased software quality prediction, *International Journal of Software Engineering and Knowledge Engineering*, 2000, Vol. 10, N 2, pp. 139–152.
- [6] Khoshgoftaar T., Seliya N. Analogy-Based Practical Classification Rules for Software Quality Estimation, *Empirical Software Engineering*, 2003, Vol. 8, N 4, pp. 325–350.
- [7] Khoshgoftaar T., Allen E.B., Deng J. Using regression trees to classify fault-prone software modules, *IEEE Transactions on Reliability*, 2002, Vol. 51, N 4, pp. 455–462.
- [8] Говорущенко Т.О. Теоретичні та прикладні засади інформаційної технології оцінювання достатності інформації щодо якості у специфікаціях вимог до програмного забезпечення – Кваліфікаційна наукова праця на правах рукопису. Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 «Інформаційні технології» (12 – Інформаційні технології) – Хмельницький національний університет, Українська академія друкарства, Львів, 2018, 310 с.
- [9] Незамай Б. С. Прогнозне моделювання показників якості програмних продуктів, *Education and science of today: intersectoral issues and development of sciences: Collection of scientific papers «ΛΟΓΟΣ» with Proceedings of the III International Scientific and Practical Conference*, Cambridge, May 20, 2022. CambridgeVinnitsia: P.C. Publishing House & European Scientific Platform, 2022, pp.180-181.
- [10] Незамай Б. С. Методика прогнозування якості програмних засобів, *Перспективи розвитку науки, освіти та технологій в контексті євроінтеграції: збірник тез доповідей міжнародної науково-практичної конференції (Полтава, 18 серпня 2022 р.)*, Полтава: ЦФЕНД, 2022, с. 65-66.

Mathematical model for studying the thermal regime of the slag-splashing lance

Oleg Zhulkovskyi, Inna Zhulkovska, Serhii Panteikov,
Kyrylo Muzychka, Olena Panteikova, Yulia Pavlenko

*Department of Systems Software; Department of Metallurgy of Ferrous Metals and Metal Forming
Dniprovsky State Technical University
Kamianske, Ukraine*

Abstract — A mathematical model has been developed to study the thermal regime of a special spray lance for applying liquid slag with increased scull properties to the unit walls to increase the productivity of steelmaking converters and increase their lining durability.

Keywords — Converter, Lining, Slag, Lance, Mathematical Model, Thermal Regime.

I. INTRODUCTION

One of the main ways to increase the steelmaking converters productivity is to increase the durability of the lining. The main factors providing an increase in the lining durability are [1] the use of special magnesite fluxes for the formation of final slag with increased scull properties and its application (spray) with the help of a special lance over the converter walls during the intermelting period.

An important problem associated with the development and implementation of these technologies is the design and increasing of the durability of special lance devices for blowing the final slag, as well as the combination of spray and flame gunning operations.

It is obvious that the actual problem, in this case, is the study of the thermal regime of the introduced blowing devices. Taking into account the multidimensionality and nonlinearity of high-temperature processes, as well as the complexity of conducting such laboratory and industrial experiments, computer modelling can be considered to be the most appropriate method for such studies. Thus, at this stage of the work, the task is to develop a mathematical model for studying the thermal regime of the spray lance barrel.

II. RELATED WORK

A review of the literature shows a sufficient variety of works devoted to the study of repairing the steelmaking converters lining process using slag (scull) lances, including the method of physical modelling, for example [2, 3]. At the same time, there is an almost complete absence of mathematical (computer) models for the thermal regime of the devices barrel under consideration, as well as the corresponding numerical studies. It should be noted that a mathematical model [4] was previously developed for the blowing top oxygen lances, and an information-modelling forecasting system for thermal regime was created on its basis [5].

III. METHODOLOGY

In the accepted problem formulation, the converter contains liquid slag from the previous melt with increased scull properties, and the bottom part of the object under study (a special top lance) is located in the converter cavity for blowing the slag onto the unit walls. The shape and axisymmetric configuration of the outer tube of the lance (barrel), as well as the heat transfer conditions, make it possible to apply a cylindrical approximation and a two-dimensional computational grid in relation to the geometric description of the device.

In the mathematical description of the model, the following assumptions were made. It is assumed that there is no slag scull or any heat-insulating coatings on the lance surface, while the end parts of the device, due to the insignificance of their sizes, can be considered conditionally heat-insulated. The mathematical model does not provide changes in the geometric parameters of the computational domain associated with deterioration or slagging of the device working surfaces during the technological process. During the entire period of blowing the slag, the lance is in a diathermic medium and is constantly cooled from the inside by a stream of water or nitrogen, depending on the design. The temperature of the walls of the converter, as well as the gases leaving the unit, are assumed to be equal to the temperature of the slag in the converter throughout the entire lining repair operation.

According to the coordinate system accepted in the problem, and also due to the absence of internal heat sources, the temperature field of the structure is described by the following differential equation of heat conductivity:

$$c\rho \frac{\partial T}{\partial \tau} = \frac{1}{r} \frac{\partial}{\partial r} \left(r\lambda \frac{\partial T}{\partial r} \right) + \frac{\partial}{\partial z} \left(\lambda \frac{\partial T}{\partial z} \right) \quad (1)$$

with boundary conditions that take into account the complex radiation-convective heat transfer of the outer barrel surface and the developed turbulent mode of convective heat transfer along the device inner surface:

$$\left(\frac{\partial T}{\partial z} \right)_t = 0; \left(\frac{\partial T}{\partial z} \right)_b = 0; -\lambda \left(\frac{\partial T}{\partial r} \right)_i = q_{\Sigma i}; -\lambda \left(\frac{\partial T}{\partial r} \right)_e = q_{\Sigma e} \quad (2)$$

In Eqs. (1) and (2): ρ is the density, kg/m³; c is the specific heat capacity, J/(kg K); T is temperature, K; τ is time, s; λ is the heat conductivity, W/(m K); q_{Σ} is total density of heat flux for the computational domain corresponding surface, W/m².

Indices in Eqs. (2): t (top), b (bottom), i (internal); e (external) designate a correspondent surface of the blowing device involved in heat exchange.

Thermal-physical quantities in the presented equations are obtained as an approximation by temperature functions of corresponding table values for pipe steel.

The initial condition of implementation of the given problem: $T(r, z, 0) = T_0$, where T_0 – temperature of the surrounding environment outside the converter (as well as the temperature of the coolant at the inlet to the lance).

Finite difference approximation of equation (1), as well as boundary conditions (2), were obtained using an integro-interpolation method (balance method), i.e. from direct approximation of relations of thermal balance denoted for elementary volumes with the fulfillment of the matching conditions for heat fluxes at the boundaries of the computational domain. The heat balance equations have been written in an implicit, split in terms of spatial variables.

The initial heat transfer equations are presented as dependencies between dimensionless similarity numbers.

The change in the temperature of water or air in the outer contour of the lance is described by the Newton-Richman law. In gas cooling, radiative heat transfer between the coaxial tubes of the considered lance circuit is also taken into account.

While calculating the radiative heat transfer of the lance barrel, the problem of heat exchange by radiation is solved for a complex system of three gray radiating surfaces: the barrel – the slag mirror – the converter walls separated by the diathermic medium of the exhaust gases. It is convenient to solve this problem by the zonal method. As a result, the problem of radiative heat transfer is reduced to finding the optical-geometric characteristics of radiation – the slope coefficients for a geometrically complex system.

The description of methods for calculating the quantities and coefficients present in equations (2) and in their finite-difference approximation, as well as changes in the water temperature values in the outer loop along the length of the lance, is described in detail in work [4]. The adequacy of the presented mathematical description is also shown there.

The general technique for numerical modelling of combined (radiative-convective) heat transfer processes in technological systems is presented in [6].

IV. RESULTS

Thus, a mathematical description of the special top lances barrel temperature regime during the operation of blowing slag onto the walls to repair the converter lining to increase its productivity has been developed.

V. DISCUSSION

As part of the research, a software functionality was developed that implements the classical right tridiagonal matrix algorithm and the two-sided Thomas algorithm parallelized into two streams. The two-sided Thomas algorithm is a combination of forward and backward sweep methods. Computer modelling is planned to be carried out using a multi-core computing architecture, which will increase the calculation time by about 2-3 times [7].

VI. FUTURE RESEARCH

Based on the obtained mathematical description, it is planned to develop a computer model, as well as an information-modelling forecasting system for thermal regime of the spray lance according to the type [5]. Such systems make it possible to evaluate the design and technological operation parameters of the blowing device as a criterion for its safe work. Their use in the "advisor" mode ensures the optimal design of lances with a rational cooling system, ensuring proper thermal conditions throughout the entire operation time, as well as long trouble-free operation.

VII. CONCLUSION

To increase the productivity of steelmaking converters and increase the lining durability by applying liquid slag with increased scull properties to the converter walls, a mathematical model has been developed to study the thermal regime of a special spray lance.

The presented mathematical description will form the basis for the development of an information-modelling forecasting system for thermal regime of the spray converter lance and effective computer modelling using modern computing architecture.

REFERENCES

- [1] *Slag splashing technique in converter operation*, Ispat Guru, March 24, 2013. [Online] Available: <https://www.ispatguru.com/slag-splashing-technique-in-converter-operation/>
- [2] О. С. Пантейкова и С. П. Пантейков, "Исследование процесса ошлаковывания рабочей поверхности футеровки 260-тонных конвертеров верхней двухъярусной гарнисажной фурмой со сгруппированными соплами Лаваля разного диаметра," в *Inżynieria i technologia. Obiecujące osiągnięcia naukowe*, Warszawa, 2017, с. 14–19.
- [3] V. Sinelnikov, M. Szucki, T. Merder, J. Pieprzyca, and D. Kalisz, "Physical and Numerical Modeling of the Slag Splashing Process," *Materials*, vol. 9, no. 14, pp. 1–19, 2021, doi: 10.3390/ma14092289.
- [4] O. A. Zhulkovskii. "Temperature conditions for operation of top-mounted bof lance barrel", *Steel in Translation*, vol. 28, no. 7, pp. 37–39, 1998.
- [5] O. A. Zhulkovskii, S. P. Panteikov, and I. I. Zhulkovskaya, "Information–Modeling Forecasting System for Thermal Mode of Top Converter Lance," *Steel in Translation*, vol. 52, no. 5, pp. 495–502, 2022, doi: 10.3103/S0967091222050138.
- [6] Жульковский О. А., Жульковская И. И., "Особенности математического моделирования процессов комбинированного теплообмена в технологических системах", *Математическое моделирование*, вып. 1, №34, с. 7-10, 2016, *Днепродзержинск, изд. ДГТУ*
- [7] O. O. Zhulkovskiy, I. I. Zhulkovska, V. V. Shevchenko, "Evaluating the effectiveness of the implementation of computational algorithms using the OpenMP standard for parallelizing programs", *Informatics and Mathematical Methods in Simulation*, vol. 11, no. 4, pp. 268-277, 2021, doi: 10.15276/imms.v11.no4.268

Системний підхід до математичного опису дифузійних процесів у випадково неоднорідних двофазних структурах з кульовими включеннями

Анастасія Чучвара¹, Ольга Чернуха^{1,2}, Юрій Білушак^{1,2}

¹ Відділ числових методів математичної фізики,
Інститут прикладних проблем механіки і математики ім. Я.С. Підстригача НАНУ
Львів, Україна

² Кафедра обчислювальної математики та програмування Інституту прикладної математики та фундаментальних наук,
Національний університет “Львівська політехніка”
Львів, Україна

Анотація—У роботі запропоновано структурну схему системного підходу до математичного опису дифузійних процесів у двофазних структурах з кульовими включеннями, розташування яких визначається різними варіантами ймовірнісних розподілів. На прикладі двофазної структури з кульовими включеннями, зосередженими біля нижньої границі тіла, наведено застосування такого підходу до дослідження масоперенесення домішкової речовини у тілі.

Ключові слова— Системний аналіз, Математична модель, Випадково неоднорідна структура, Кульові включення, Бета-розподіл, Ряд Неймана.

I. ВСТУП

Структури з внутрішніми неоднорідностями у вигляді пор чи гранул знайшли широке використання у будівництві (газо-, пінобетон), медицині (титанові композити з кульовими магнієвими включеннями), сучасному авіа- та машинобудуванні (композитні матеріали на основі алюмінію, карбону), тощо [1]-[3]. Наявність внутрішніх неоднорідностей у структурі значно впливає на хімічну та фізичну взаємодію твердих речовин з газами та рідинами, однак проведення натурних експериментів для прогнозування та визначення поведінкових властивостей таких структур є часо- та ресурсозатратним процесом. А отже, актуальним є створення математичних моделей та проведення на їх основі комп'ютерного моделювання процесів перенесення маси у неоднорідних структурах з метою встановлення особливостей матеріалів, їх міцності, надійності, тощо.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Відповідно до характеристик, які вимагається дослідити, існують різні підходи до опису дифузійних процесів у неоднорідних, зокрема, пористих структурах. У випадку, якщо відоме поле швидкостей, для дослідження явища дифузії простим у реалізації є метод випадкового блукання [4], [5]. В останні роки широкого використання набув метод решіток Больцмана. У праці [6] цей метод застосовано для розв'язання рівняння дифузії у пористому середовищі, змодельованому двовимірними діаграмами Вороного. У дослідженні [7] для відслідковування руху частинок у тілі зі структурними неоднорідностями, згенерованими методами Монте-Карло та Джорді-Торі, використано комбінацію методів решіток Больцмана та випадкового блукання. Поряд із багатьма перевагами методу решіток Больцмана варто від-

значити, що його реалізація вимагає значних обчислювальних потужностей, а використання на рухомих границях та для багатозафазних чи багатоконпонентних моделей є ускладненим.

Водночас за рахунок неповної інформації щодо внутрішньої структури системи виникають умови невизначеності, що полягають у відсутності інформації про точне розташування включень, їх кількість, тощо, однак задають розподіл вірогідності на безліч результатів. Тоді така система розглядається як стохастично невизначена.

Недостатність інформації про внутрішню структуру об'єкта можна подолати застосуванням методів гомогенізації, що ставлять у відповідність стохастичній системі детерміновану з ефективними коефіцієнтами [8]. Проте методи гомогенізації потребують жорстких обмежень на параметри системи, включно з величиною часів релаксації процесів, розмірами неоднорідностей тіла і самого об'єкта, рівномірним розподілом структурних особливостей і/або елементів, тощо [9]. Невиконання хоча б однієї з умов застосовності гомогенізації суттєво збільшує ризики отримання недостовірного результату. Тому необхідно розробляти нові підходи і методи, будувати нові математичні моделі, які б враховували вплив умов невизначеності системи.

У цій роботі запропоновано структурну схему для якісного та кількісного опису процесів масоперенесення в двофазних тілах з випадково розташованими кульовими включеннями, на основі підходу, відповідно до якого крайові задачі дифузії зводяться до еквівалентного інтегро-диференціального рівняння, розв'язок якого будують у вигляді ряду Неймана із подальшим його усередненням за ансамблем конфігурацій фаз [10], [11].

III. СТРУКТУРНА СХЕМА ДОСЛІДЖЕННЯ ДИФУЗІЙНИХ ПРОЦЕСІВ У ВИПАДКОВО НЕОДНОРІДНИХ СТРУКТУРАХ

Приймаючи, що відсутність точної геометричної конфігурації внутрішньої структури тіла є умовою невизначеності для даної системи, побудовано структурну схему системного аналізу для математичного опису процесів дифузії в тілі з випадково розташованими кульовими включеннями.

Конкретизацію підходу проведено для випадку, показаного на рис.1, коли кульові включення найбільш ймовірно розташовані біля нижньої границі тіла.

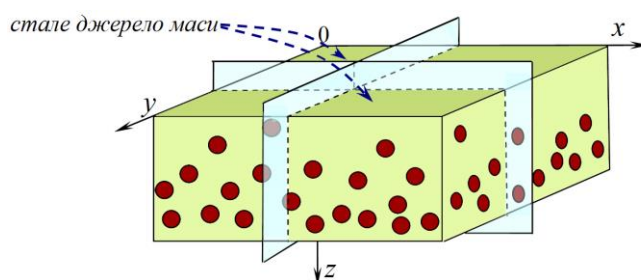


Рис. 1. Схематичне представлення випадкової структури двофазного тіла з кульовими включеннями, сконцентрованими біля нижньої границі тіла

Прийнято, що вздовж осей Ox та Oy реалізується рівномірний розподіл фаз, а вздовж осі Oz - частковий випадок бета-розподілу з густиною ймовірностей [12]

$$f(z) = \begin{cases} \alpha \left(\frac{z}{z_0} \right)^{\alpha-1}, & z \in [0; z_0] \\ 0, & z \notin [0; z_0] \end{cases} \quad (1)$$

де α - ступінь вільності розподілу ($\alpha > 1$), z_0 - товщина шару.

Запропонований підхід містить одинадцять етапів, частина з них поділяється на підетапи. Структурна схема такого підходу представлена на рис. 2.

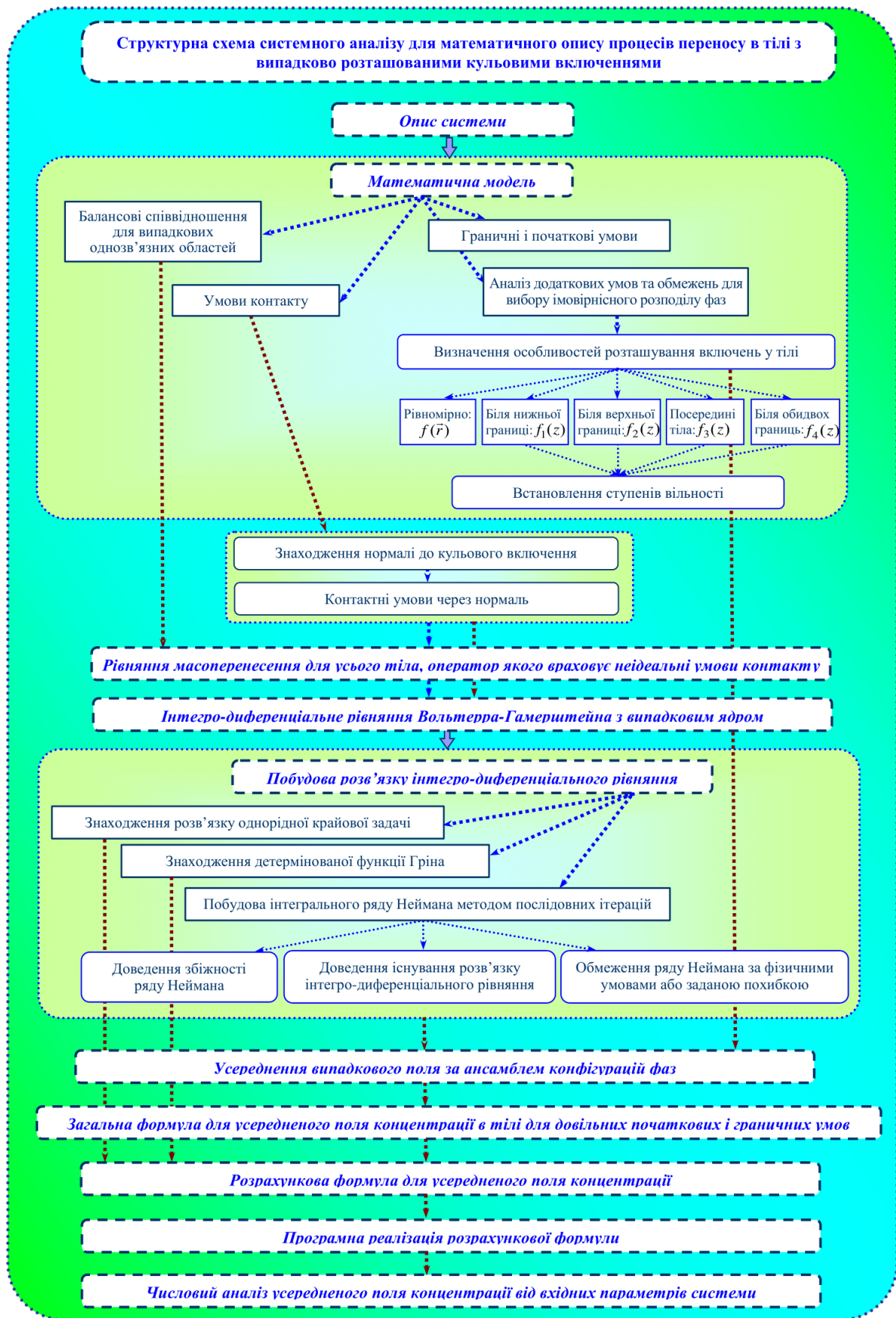


Рис. 2. Структурна схема системного підходу до математичного опису процесів дифузії у двофазних структурах з кульовими включеннями

Еман I. Опис системи включає опис об'єкта дослідження та його структуру, процеси, які протікають у цьому об'єкті, різного типу умови, в яких перебуває система, а також обмеження, які необхідно накладати як на самий об'єкт дослідження, так і на процеси, і, за необхідності, на систему в цілому [13].

В якості конкретизації розглянемо дифузію домішкової речовини у двофазному тілі з випадково розташованими кульовими включеннями у випадку, коли необхідно явно враховувати величину стрибків шуканої функції та її похідної на міжфазних границях. Включення найбільш ймовірно знаходяться біля нижньої границі тіла (рис.1). Приймаємо, що об'ємна частка матриці є набагато більшою, ніж об'ємна частка включення: $v_0 \gg v_1$, проте умови на розміри включень не накладаються. Густина тіла $\rho(\vec{r})$ та коефіцієнт дифузії $d(\vec{r})$ є сталими в об'ємі кожної фази ($\vec{r} = (x, y, z)$ - радіус-вектор біжучої точки).

Еман II. На основі фізичного опису системи, встановлення загальних необхідних умов та обмежень будемо математичну модель. При цьому можна виділити такі підетапи:

Підетап 1. Записуємо балансові співвідношення для (фізичних, хімічних тощо) процесів, які протікають в тілі в цілому або в кожному структурному елементі в залежності від типу об'єкта і мети дослідження [14].

Для процесу дифузії у двофазній смузї з випадковими включеннями складаємо балансові співвідношення процесу масоперенесення для всіх однозв'язних областей кожної з фаз. З урахуванням I та II законів Фіка отримаємо [11]

$$\rho_j \frac{\partial c_j(\vec{r}, t)}{\partial t} = d_j \Delta c_j(\vec{r}, t), \quad \vec{r} \in (V^{(j)}), \quad t \in [0; \tau] \quad (\tau < \infty),$$

де ρ_j - густина випадкової області $(V^{(j)})$; d_j - кінетичний коефіцієнт переносу в цій області (індекс $j = 0$ відповідає матриці, $j = 1$ - включенням).

Підетап 2. Записуємо умови контакту між різними фазами.

Приймаємо, що на випадкових границях розділу областей $(V^{(0)})$ і $(V^{(1)})$ виконуються умови рівності хімічних потенціалів та дифузійних потоків частинок. Ці умови зводяться до неідеальних контактних умов на функцію концентрації [15].

Підетап 3. Формулюємо початкові та граничні умови:

$$c_0(\vec{r}, t)|_{t=0} = c_1(\vec{r}, t)|_{t=0} = 0, \quad c_0(\vec{r}, t)|_{z=0} = c_* \equiv const, \quad c_0(\vec{r}, t)|_{z=z_0} = 0,$$

$$c_j(\vec{r}, t)|_{x, y \rightarrow \pm\infty} \leq K < \infty \quad (j = \overline{0,1}).$$

Підетап 4. Проводимо аналіз додаткових умов та обмежень для вибору імовірнісного розподілу фаз.

Підетап 5. Визначаємо закономірності розташування включень в області тіла та відповідний ймовірнісний розподіл [12]:

- фази розподілені рівномірно, густина функції розподілу $f(\vec{r}) = V_{phase}/V_{body}$;
- включення найімовірніше знаходяться біля нижньої границі тіла (частковий випадок

$$\text{бета-розподілу), } f_1(z) = \begin{cases} \alpha \left(\frac{z}{z_0} \right)^{\alpha-1}, & z \in [0; z_0] \\ 0, & z \notin [0; z_0] \end{cases}$$

- включення найімовірніше знаходяться біля верхньої границі тіла (частковий випадок бета-розподілу), $f_2(z) = \begin{cases} (1 + \beta) \left(1 - \frac{z}{z_0}\right)^{\beta-1}, & z \in [0; z_0]; \\ 0, & z \notin [0; z_0], \end{cases}$
- включення найімовірніше знаходяться посередині тіла (бета-розподіл), $f_3(z) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{z}{z_0}\right)^{\alpha-1} \left(1 - \frac{z}{z_0}\right)^{\beta-1}, & z \in [0; z_0]; \\ 0, & z \notin [0; z_0] \end{cases}$
- включення найімовірніше знаходяться біля обох границь тіла (arcsin-розподіл), $f_4(z) = \pi^{-1} \left[\frac{z}{z_0} \left(1 - \frac{z}{z_0}\right) \right]^{-1/2}$, тощо.

Підман 6. За необхідності з експериментальних даних встановлюємо ступені вільності розподілів α, β .

Таким чином, математична модель сформульована.

Оскільки ми сформулювали диференціальні рівняння дифузії для кожної однозв'язної області, то при «зшивці» областей потрібно враховувати нормаль, яка є зовнішньою для даної області. Тому наступний етап присвячений явному врахуванню нормалі в умовах контакту.

Еман III. Врахування нормалі до кульового включення:

Підман 1. Знаходимо рівняння нормалі [16] до кульового включення в деякій точці поверхні \vec{r}_{Γ_i} :

$$\vec{n}_i = \frac{x - x_{\Gamma_i}}{R_j(x_{\Gamma_i} - x_{c_i})} \vec{i} + \frac{y - y_{\Gamma_i}}{R_j(y_{\Gamma_i} - y_{c_i})} \vec{j} + \frac{z - z_{\Gamma_i}}{R_j(z_{\Gamma_i} - z_{c_i})} \vec{k},$$

де $(x_{c_i}, y_{c_i}, z_{c_i}) = \vec{r}_{c_i}$ - центр i -го кульового включення, R_j - характерний радіус включення j -ої фази.

Підман 2. Записуємо неідеальні умови контакту через нормаль

$$[c(\vec{r}, t)]_{\vec{r} \in \Gamma} = \left(1 - \frac{k_0}{k_1}\right) c_0(\vec{r}, t)|_{\vec{r} \in \Gamma-0} \vec{n}, \quad [\vec{\nabla} c(\vec{r}, t)]_{\vec{r} \in \Gamma} = \left(1 - \frac{\rho_0 d_0}{\rho_1 d_1}\right) \vec{\nabla} c_0(\vec{r}, t)|_{\vec{r} \in \Gamma-0} \vec{n}.$$

Тут $[c(\vec{r}, t)]_{\vec{r} \in \Gamma}$ - стрибок функції $c(\vec{r}, t)$ на границях контакту Γ , k_j ($j = 0; 1$) - кінетичні коефіцієнти, що визначають величину стрибків $c(\vec{r}, t)$ на Γ .

Еман IV. Отримання рівняння масоперенесення для усього тіла, оператор якого враховує неідеальні умови контакту [11].

Вводячи в розгляд випадкову функцію просторової координати $c(\vec{r}, t)$, яка описує концентрацію в усьому тілі, і враховуючи, що вона та її похідна мають розриви 1-го роду на границях контакту областей, із застосуванням теорії узагальнених функцій отримаємо

$$\rho(\vec{r}) \frac{\partial c(\vec{r}, t)}{\partial t} = \{d(\vec{r})\} \{\Delta c(\vec{r}, t)\} + [d(\vec{r})]_{\Gamma} [\vec{\nabla} c(\vec{r}, t)]_{\Gamma} \delta(\vec{r} - \vec{r}_{\Gamma}) +$$

$$+ [d(\vec{r})]_{\Gamma} [c(\vec{r}, t)]_{\Gamma} \vec{\nabla} \delta(\vec{r} - \vec{r}_{\Gamma}) + [d(\vec{r})]_{\Gamma} [c(\vec{r}, t)]_{\Gamma} \delta^2(\vec{r} - \vec{r}_{\Gamma}),$$

де \vec{r}_{Γ} – радіус-вектор точок границі Γ , $\{\dots\}$ – області неперервності функції, $[\dots]_{\Gamma}$ – стрибок функції на границі Γ , $\delta(\cdot)$ – дельта-функція Дірака.

Еман V. Отримання еквівалентного вихідній контактної-крайовій задачі інтегро-диференціального рівняння з випадковим ядром.

Розглядаючи неоднорідність структури тіла як внутрішні джерела, будемо інтегро-диференціальне рівняння як суму розв'язку однорідної крайової задачі та згортки функції Гріна з джерелом

$$c(\vec{r}, t) = c_0(\vec{r}, t) + \int_0^t \iiint_V G(\vec{r}, \vec{r}', t, t') L_s(\vec{r}', t') c(\vec{r}', t') d\vec{r}' dt'. \quad (2)$$

Тут $c_0(\vec{r}, t)$ – розв'язок однорідної крайової задачі [11], $G(\vec{r}, \vec{r}', t, t')$ – детермінована функція Гріна [10], $L_s(\vec{r}', t')$ – випадковий оператор вигляду

$$\begin{aligned} L_s(\vec{r}, t) = & (\rho_0 - \rho_1) \sum_{i=1}^{n_1} \eta_{i1}(\vec{r}) \frac{\partial}{\partial t} - (d_0 - d_1) \sum_{i=1}^{n_1} \eta_{i1}(\vec{r}) \Delta + \\ & + |d_0 - d_1| \sum_{i=1}^{n_1} \vec{n}_i \cdot \vec{n}_i \left(1 - \frac{\rho_0 d_0}{\rho_1 d_1} \right) \vec{\nabla} \dots \Big|_{\vec{r} \in \Gamma_{i1}-0} \delta(\vec{r} - \vec{r}_{\Gamma_{i1}}) + \\ & + |d_0 - d_1| \sum_{i=1}^{n_1} \vec{n}_i \cdot \vec{n}_i \left(1 - \frac{\rho_0 d_0}{\rho_1 d_1} \right) \vec{\nabla} \dots \Big|_{\vec{r} \in \Gamma_{i1}-0} \vec{\nabla} \delta(\vec{r} - \vec{r}_{\Gamma_{i1}}) + \\ & + |d_0 - d_1| \sum_{i=1}^{n_1} \vec{n}_i \cdot \vec{n}_i \left(1 - \frac{\rho_0 d_0}{\rho_1 d_1} \right) \vec{\nabla} \dots \Big|_{\vec{r} \in \Gamma_{i1}-0} \delta^2(\vec{r} - \vec{r}_{\Gamma_{i1}}). \end{aligned}$$

Еман VI. Побудова розв'язку отриманого інтегро-диференціального рівняння.

Розв'язок інтегро-диференціального рівняння (2) шукаємо у вигляді інтегрального ряду Неймана методом послідовних ітерацій. Нехтуючи парним взаємовпливом включень, обмежуємося двома першими членами ряду Неймана.

Еман VII. Усереднення випадкового поля, знайденого у вигляді ряду Неймана, за ансамблем конфігурацій фаз.

Вважаємо, що випадковою величиною, за якою проводимо процедуру усереднення, є радіус-вектор центра кожного кульового включення \vec{r}_{i1} . Враховуємо, що функції $c_0(z, t)$ і $G(\vec{r}, \vec{r}', t, t')$ є детермінованими, а випадковою є функція структури $\eta_{i1}(\vec{r})$. Тоді, враховуючи рівномірний розподіл включень вздовж осей Ox та Oy , а також частковий випадок бета-розподілу (1) вздовж осі Oz , за теоремами теорії ймовірностей усереднення за ансамблем конфігурацій фаз зводиться до знаходження такого інтегралу

$$I_{\eta} = \frac{1}{\sqrt{V}} \sum_{i=1}^{n_1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{R_1}^{z_0 - R_1} \eta_{i1}(|\vec{r}' - \vec{r}_{i1}|) f_2(z_{i1}) dz_{i1} dy_{i1} dx_{i1},$$

де \tilde{V} – розміри тіла за осями Ox та Oy , R_1 – характерний радіус сферичних включень.

Eman VIII. Отримання для довільних початкових і граничних умов загальної формули для усередненого поля концентрації за рівномірного розподілу включень вздовж осей Ox та Oy і бета-розподілу (1) вздовж осі Oz .

Для усередненого поля концентрації домішкової речовини у двофазній смузі зі структурою, поданою на рис.1, отримана наступна загальна формула

$$\begin{aligned} \langle c(\vec{r}, t) \rangle_{conf} = & c_0(z, t) + \frac{3\alpha v_1}{2z_0^{\alpha-2}} \int_0^{+\infty} \int_{-\infty}^{+\infty} \sum_{k=0}^{\alpha-1} \frac{(1+(-1)^k)(z')^{\alpha-k-1}(\alpha-1)!}{(k+1)!(\alpha-k-1)!(k+3)} \times \\ & \times R_1^k \left(\int_{R_1}^{2R_1} \left(\frac{z'}{R_1} - 1 \right)^{k+3} G(\vec{r}, \vec{r}', t, t') \left[(\rho_0 - \rho_1) \frac{\partial c_0}{\partial t'} - (d_0 - d_1) \frac{\partial^2 c_0}{\partial z'^2} \right] dz' + \right. \\ & \left. + \int_{2R_1}^{z_0} G(\vec{r}, \vec{r}', t, t') \left[(\rho_0 - \rho_1) \frac{\partial c_0}{\partial t'} - (d_0 - d_1) \frac{\partial^2 c_0}{\partial z'^2} \right] dz' \right) dy' dx' dt'. \end{aligned} \quad (3)$$

Eman IX. Знаходження розрахункової формули для усередненого поля концентрації через підстановку в загальну формулу (3) виразів для концентрації домішки в однорідному тілі та функції Гріна для конкретних початкових і граничних умов, які задані у вихідній постановці задачі або відповідають їм.

Отримано розрахункову формулу для усередненого за ансамблем конфігурацій фаз поля концентрації за часткового випадку бета-розподілу включень, що відповідає структурі, у якій область найбільш ймовірного розташування кульових включень зосереджена біля нижньої границі тіла, у вигляді

$$\begin{aligned} \langle c(\vec{r}, t) \rangle_{conf} \approx & c_0(z, t) + \alpha \frac{4c_* \tilde{v}_1 \pi K_{dp}}{z_0^{\alpha-1}} \sum_{l=1}^{\infty} \sin(y_l z) \sum_{k=0}^{\alpha-1} F(\alpha, k) R_1^{k-1} \times \\ & \times \left[y_l \tilde{A}_k e^{\frac{d_0 y_l^2 t}{\rho_0}} + \sum_{\substack{n=1 \\ n \neq l}}^{\infty} A_k^{ln} \left(e^{\frac{d_0 y_n^2 t}{\rho_0}} - e^{\frac{d_0 y_l^2 t}{\rho_0}} \right) \right], \end{aligned} \quad (4)$$

$$\text{де } K_{dp} = \frac{d_0}{\rho_0} \left(\frac{d_1}{d_0} - \frac{\rho_1}{\rho_0} \right); \quad F(\alpha, k) = \frac{(\alpha-1)!(1+(-1)^k)}{(k+1)!(\alpha-k-1)!(k+3)};$$

$$A_k^{ln} = \frac{y_n \rho_0}{d_0 (y_l^2 - y_n^2)} \left(\frac{I_{11} - I_{12}}{R_1^{k+3}} + I_{21} - I_{22} \right), \quad \tilde{A}_k = \frac{\tilde{I}_{11} - \tilde{I}_{12}}{R_1^{k+3}} + \tilde{I}_{21} - \tilde{I}_{22}, \quad y_{ln}^{\pm} = y_l \pm y_n,$$

$$I_{11} = I(R_1, 2R_1; y_{ln}^-, k), \quad I_{12} = I(R_1, 2R_1; y_{ln}^+, k), \quad I_{21} = I(2R_1, z_0; y_{ln}^-, k), \quad I_{22} = I(2R_1, z_0; y_{ln}^+, k),$$

$$\tilde{I}_{11} = I_0(R_1, 2R_1; k), \quad \tilde{I}_{12} = I(R_1, 2R_1; 2y_l, k), \quad \tilde{I}_{21} = I_0(2R_1, z_0; k), \quad \tilde{I}_{22} = I(2R_1, z_0; 2y_l, k).$$

Eman X. На основі розрахункової формули (4) створення відповідного програмного забезпечення.

Розроблений комплекс програм включає програми розрахунку усередненого за ансамблем конфігурацій фаз поля концентрації за часткового випадку бета-розподілу включень для найбільш ймовірного розташування кульових включень біля нижньої границі тіла; детермінованої функції Гріна; концентрації частинок в однорідному тілі.

Етап XI. Проведення числового аналізу усередненого поля концентрації розвиненими програмними засобами.

Для кількісного дослідження масоперенесення домішкової речовини у двофазному тілі за умов наймовірнішого розташування сферичних включень біля нижньої границі тіла, використовуємо розрахункову формулу (4). Обчислення виконані у безрозмірних змінних $\zeta = z/z_0$ та $\tau = d_0 t / z_0^2$. В якості базових параметрів прийнято наступні значення коефіцієнтів: $\tau = 0,4$; $d_1/d_0 = 0,01$ (криві а) та $d_1/d_0 = 5$ (криві б); $v_1 = 0,2$, $R_1 = 0,001$; $\rho_1/\rho_0 = 1,3$.

На рис. 3 наведено розподіли усередненої концентрації домішкової речовини у двофазному тілі зі сферичними включеннями, структура якого подана на рис.1 для різних значень об'ємної частки включень $v_1 = 0,2$; $0,25$; $0,3$ (криві 1-3, рис. 3а) та ступенів вільності $\alpha = 3$; 5 ; 7 (криві 1-3, рис. 3б).

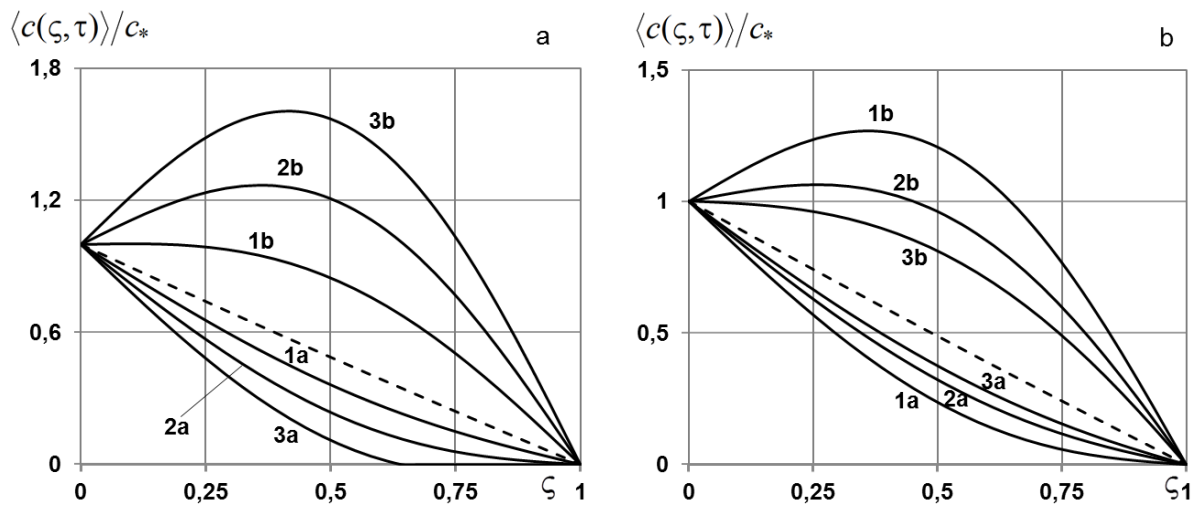


Рис. 3. Розподіли усередненої концентрації домішкової речовини для різних значень об'ємної частки включень (а) та ступенів вільності бета-розподілу (б)

Збільшення об'ємної частки включень для випадку $d_1 < d_0$ зменшує значення усередненої концентрації домішкової речовини у неоднорідній структурі (криві а, рис. 3а), водночас для випадку $d_0 < d_1$ збільшення v_1 викликає зростання усередненої концентрації і значення $\langle c(\zeta, \tau) \rangle / c_*$ в середині тіла можуть стати більшими, ніж значення концентрації на верхній границі тіла (криві 2b та 3b, рис. 3а). Зазначимо, що відповідно до властивостей функції щільності бета-розподілу, збільшення ступеня вільності розподілу α відповідає ущільненню кульових включень до нижньої поверхні тіла, протилежної до джерела маси. Для випадку $d_1/d_0 < 1$ ущільнення кульових включень до поверхні тіла призводить до збільшення усередненої концентрації домішки у неоднорідному тілі, та наближенні функції $\langle c(\zeta, \tau) \rangle / c_*$ до концентрації домішкової речовини в однорідному тілі (криві а, рис. 3б).

IV. ВИСНОВКИ

Таким чином, розроблено структурну схему системного підходу, яка включає побудову математичної моделі і розвинення методології розв'язування відповідних контактнокрайових задач, для якісного і кількісного опису складних систем з випадковою внутрішньою структурою. Системний підхід включає 11 послідовних етапів, починаючи від етапу «опис об'єкта», в рамках якого визначається сам об'єкт, процеси, які протікають у цьому об'єкті, різного типу умови, в яких перебуває система, а також обмеження, які необхідно накладати на систему, визначається, чи перебуває система в умовах невизначеності і який тип такої невизначеності. Запропонована схема дослідження

реалізована для математичного опису дифузійних процесів у двофазній смузі з випадково розташованими сферичними включеннями, які найімовірніше знаходяться біля нижньої поверхні тіла. Розв'язок контактної-крайової задачі представлено у вигляді інтегрального ряду Неймана та одержано загальну (для довільних початкових та граничних умов) та розрахункову формули для усередненого за ансамблем конфігурацій фаз поля концентрації частинок. На основі отриманої розрахункової формули проведено комп'ютерне моделювання та встановлено основні закономірності розподілів усередненої концентрації домішкової речовини у смузі залежно від значень коефіцієнтів дифузії, густини та об'ємної частки включень.

ЛІТЕРАТУРА

- [1] S. Jiang, L.J. Huang, Q. An, L. Geng, X.J. Wang, and S. Wang, "Study on titanium-magnesium composites with bicontinuous structure fabricated by powder metallurgy and ultrasonic infiltration", *J. of the Mech. Behav. of Biomed. Mat.*, vol. 81, pp. 10-15, 2018.
- [2] T. E. Balaji, H. Tanaya Das, and T. Maiyalagan, "Recent trends in bimetallic oxides and their composites as electrode materials for supercapacitor applications", *ChemElectroChem*, vol. 8, no. 10, pp. 1723-1746, 2021.
- [3] V. Kovbashyn and I. Bochar, "The study of technologies to improve physical-mechanical and chemical properties of reaction sintered ceramic materials on the basis of silicon carbide", *Sci. J. of TNTU*, vol. 86, no 2, pp. 14-20, 2017.
- [4] C. Vamoş, N. Suciua and H. Vereecken, "Generalized random walk algorithm for the numerical modeling of complex diffusion processes", *J. of Comput. Phys.*, vol. 186, no. 2, pp. 527-544, 2003.
- [5] E.M. LaBolle, J. Quastel, E.G. Fogg and J. Gravne, "Diffusion processes in composite porous media and their numerical integration by random walks: Generalized stochastic differential equations with discontinuous coefficients", *Water Res. Research*, vol. 36, no. 3, pp. 651-662, 2003.
- [6] Y. Yong, X. Lou, S. Li, C. Yang and X. Yin, "Direct simulation of the influence of the pore structure on the diffusion process in porous media", *Comp. & Math. with Appl.*, vol. 67, no. 2, pp. 412-423, 2014.
- [7] V. Hlushkou, S. Khirevich, V.V. Apanasovich and U. Tallarek, "Pore-scale dispersion in electrokinetic flow through a random sphere packing", *Analyt. Chem.*, no. 79, pp. 113-121, 2007.
- [8] B. Gambin, J. Telega and L. Nazarenko "Stationary thermoelasticity and stochastic homogenization", *Arch. of Mech.*, vol. 54(4), pp.319-345, 2002.
- [9] Л. Хорошун, "Математические модели и методы механики стохастических композитных материалов", *Прикл. Мех.*, Т. 36, no. 10, с. 30-62, 2000.
- [10] Є.Я. Чапля and О.Ю. Чернуха, *Фізико-математичне моделювання дифузійних процесів у випадкових і регулярних структурах*, Київ: Наукова думка, 2009.
- [11] O. Chernukha and A. Chuchvara "Modeling of the diffusion of admixtures in a porous body with random spherical pores with comparable volume fractions of the phases", *J. of Math. Sci.*, vol. 258, no. 4, pp. 553-567, 2021.
- [12] В. Корольок, Н. Портенко, А. Скороход and А. Турбин, *Справочник по теории вероятности и математической статистике*, Москва: Наука, 1985.
- [13] А.А. Самарский and А.П. Михайлов, *Математическое моделирование: Идеи. Методы. Примеры*, Москва: Физматлит, 2001.
- [14] Я. Бурак, Я. Чапля and О. Чернуха, *Континуально-термодинамічні моделі механіки твердих розчинів*, Київ: Наукова думка, 2006.
- [15] А. Мюнстер, *Химическая термодинамика*, Москва: Мир, 1971.
- [16] T. Willmore, *An introduction to differential geometry*, Courier Corporation, 2013.

On the Characteristics of the Input Material Flow of the Transport Conveyor

Oleh Pihnastyi, Maksym Sobol, Dmytro Yelchaninov

National Technical University "Kharkiv Polytechnic Institute"
Kharkiv, Ukraine

Abstract— In this paper, the statistical characteristics of the flow of material entering the input of a conveyor-type transport system are studied. For a set of data obtained as a result of experimental measurements of the input flow of material, the law of distribution of a random variable and the correlation function is investigated. Theoretical assumptions about the law of change of the correlation function for the input flow of material are confirmed.

Keywords—*Transport Conveyor, Incoming Flow Model, Data Set, Stochastic Process, Neural Network.*

I. INTRODUCTION

Belt conveyors are widely used in the mining industry, as they have relatively specific low costs for transporting rock [1]. Despite its apparent simplicity, a transport conveyor is a complex dynamic distributed system Fig.1, Fig.2. The operation of such systems is associated with high energy costs. The total length of the modern transport system exceeds 100 km, and the length of a separate section is up to 20 km. The main characteristics of modern conveyor-type transport systems are presented in [2, 3, 4].

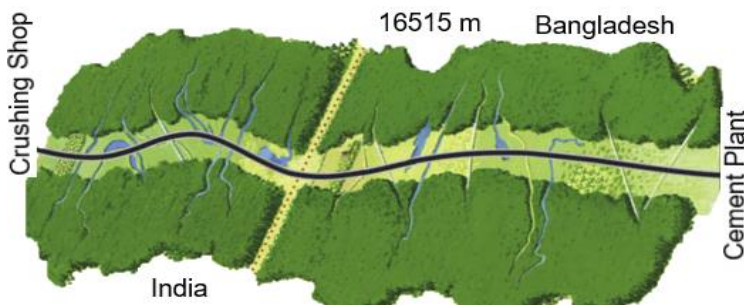


Fig.1. The long conveyor from India to Bangladesh (16.5 km)



Fig. 2. The longest belt conveyor system (Western Sahara), more than 100 km conveying length (11 flights of up to 11.7 km) [2]

Optimal control of the flow parameters of the transport system makes it possible to reduce specific transport costs by up to 30% [5]. To reduce specific transportation costs, belt speed control systems [6, 7], material consumption at the output of the bunker [8, 9], and energy management methodology [10, 11] are used. In this regard, the further development and synthesis of new algorithms for optimal control of the flow parameters of a conveyor-type transport system is an actual problem [12]. Of particular interest to researchers of transport systems are models based on neural networks [13, 14, 15]. One of the key issues in building such models is the availability of data sets for neural network training.

II. RELATED WORK

The main models that are usually used to design optimal control systems for the transport conveyor flow parameters are numerical models. The foundation of these models is the finite element method, finite difference method, Lagrange method, system dynamics method. A detailed review of the conveyor systems models shows that these models are used to determine the dynamics of changes in the individual conveyor section flow parameters [5]. It is necessary to note that the use of numerical models for the development of algorithms for optimal control of the transport system flow parameters causes significant difficulties and requires significant computational resources. The analytical PiKh-model [16] is a successful tool for describing transport systems consisting of dozens conveyor sections. The model of an individual section consists, as a rule, of two basic equations. Projected, a transport system which contains one hundred separate sections is described by a model of two hundred equations. As a result, the analytical model requires significant computational, and most importantly, time resources. The reaction time of the control system, which is based on an analytical model for such number of sections, may exceed the maximum allowable. Therefore, the analytical model did not solve the problems that occurred when using numerical models but moved the applicability of the model to a much larger number of conveyor sections.

Along with the numerical models considered above, there are many studies devoted to the use of neural networks [13, 14]. The studies demonstrate that the most important advantage of control systems based on models using a neural network is that the response time of a control system is much smaller than for numerical models. With a significant increase in the number of sections, the duration of the calculation of the output parameters can be significantly shorter than for a conveyor system analytical multi-section model. The model using a neural network can be successfully applied to describe multi-section conveyor systems if there are data to train the neural network.

III. METHODOLOGY

To build a transport system model based on a neural network, data sets for training the neural network is needed. A set of data in the form of a set of input and output flow parameters can be obtained for functioning conveyor systems. However, this set of data corresponds to a rather narrow mode of operating parameters of the production system, and it can be used to improve the existing transport conveyor control system and is not suitable for designing control systems for new transport systems. This circumstance is connected not only with the presence of different operating modes for the conveyor systems operation, but also with the structure conveyor system: the sections amount and the interactions between them. In addition to that, in many cases, collecting data for training is not possible.

One of the ways to solve this problem is to build input flow generators, which are based on the analysis of experimental data. The data sets required for training the neural network must be built in accordance with a certain law of distribution of the input material flow value. If this condition is not met, the use of these data sets for training a neural network is inefficient. For that, there is a necessity to support or decline some assumptions made in publications, in particular, assumptions about correlation function and the distribution law of the input material flow of working transport conveyor.

The statistical data of the input material flow was taken for the bucket-wheel excavator SRs 2000 [17] (Fig. 3, Fig. 4).

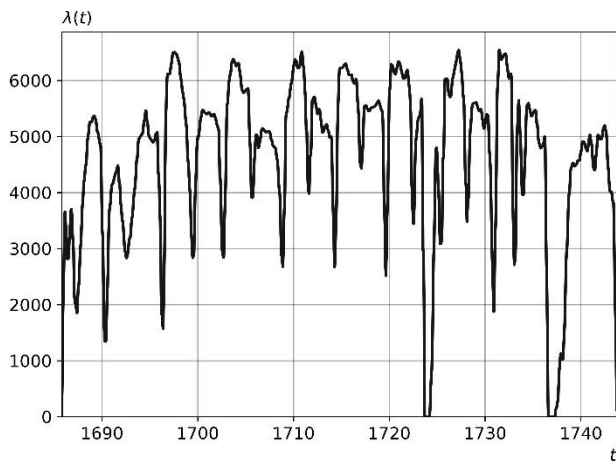


Fig. 3. Input material flow for the SRs 2000 bucket-wheel excavator [17]

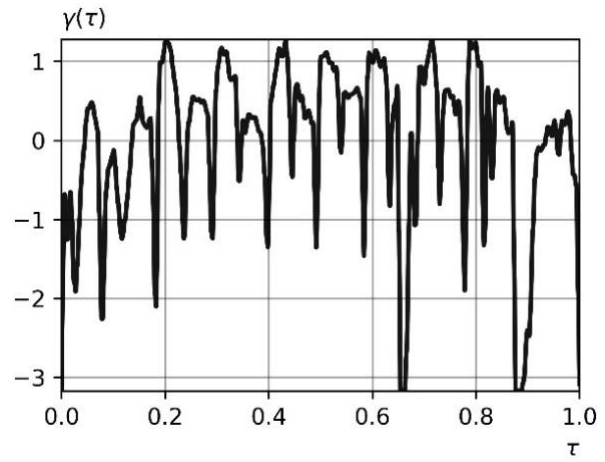


Fig. 4. Dimensionless input material flow for the SRs 2000 bucket-wheel excavator

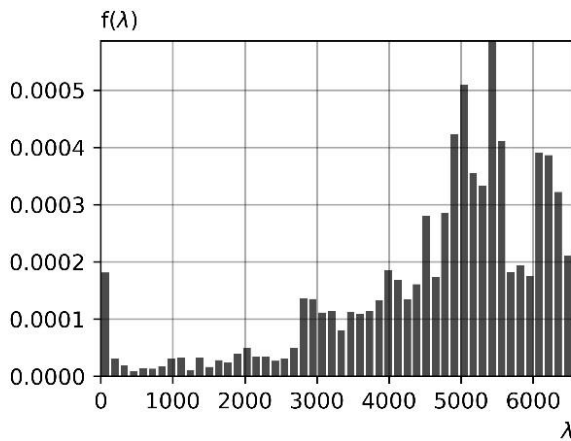


Fig. 5. Capacity distribution density for the SRs 2000 bucket-wheel excavator

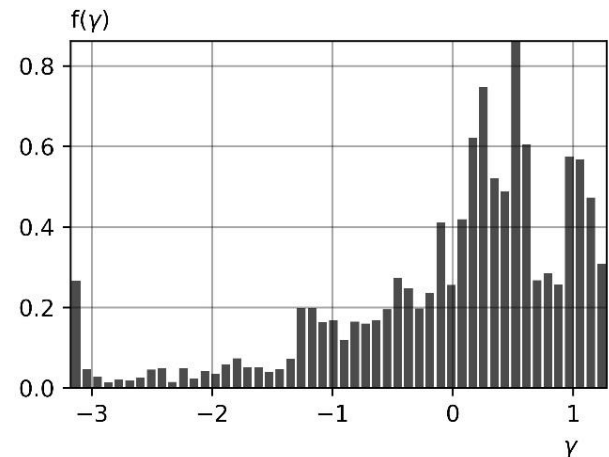


Fig. 6. Dimensionless capacity distribution density for the SRs 2000 bucket-wheel excavator

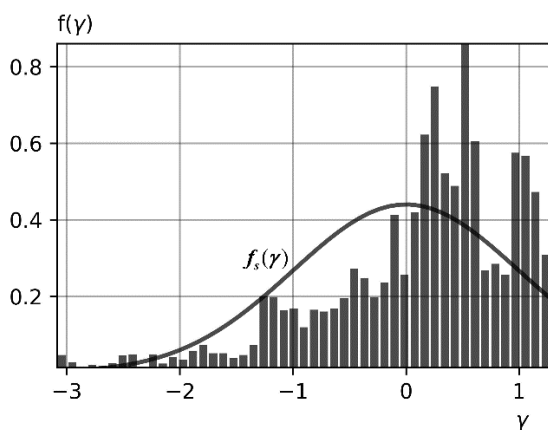


Fig. 7. Statistical and theoretical distribution of the input material flow

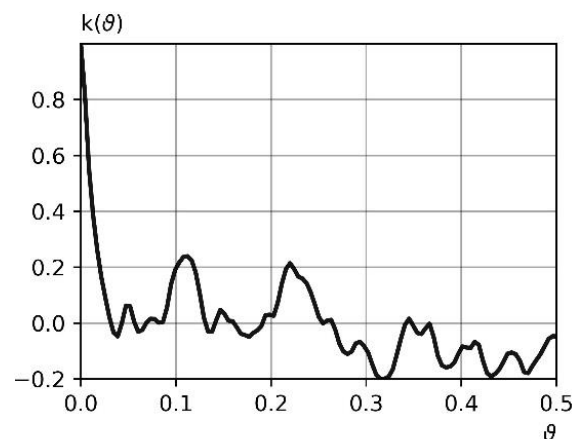


Fig. 8. Correlation function

There is an assumption that the transport conveyor input material flow random variable has a normal distribution law [18, 19, 20]. In terms of correlation function there are recommendations

that the correlation functions of the material incoming input flow value can be approximately represented in exponential form. To test these assumptions histograms for an input material flow random variable were construct (Fig. 5, Fig. 6).

Obtained in this way statistical distribution of the transport conveyor input material flow value was compared with theoretical distribution using Pearson's goodness-of-fit test χ^2 .

IV. RESULTS

Comparative analysis of the discrepancy between the statistical and theoretical distribution of the input material flow does not confirm the assumption that the transport conveyor input material flow random variable has a normal distribution law. On the other hand, the correlation function (Fig.8) has a dependence that is quite closely characterized by the theoretical correlation function. Thus, recommendations that the correlation functions of the material incoming input flow value can be approximately represented in exponential form are confirmed.

V. DISCUSSION

As a result, the implementation of a random process is analyzed. Theoretical assumptions were tested using experimental data. These results can be used to construct a generator of the input material flow based on production data. A transport system model based on a neural network trained by generated in this way data sets can be used in the synthesis of algorithms for optimal control of the conveyor type transport systems flow parameters. Additional studies are required of works that present experimental data on the magnitude of the input material flow for modern transport conveyors.

VI. FUTURE RESEARCH

The present work is the first step in such experimental data multifaceted study. Additionally, to improve the accuracy of the approximation, it is required to analyze the input material flow represented by a non-stationary random process, which is considered as a prospect for further research. An additional problem is to determine the law of distribution of the value of the input cargo material flow entering the transport system input per time unit.

VII. CONCLUSION

To reduce the specific energy consumption of material transportation using conveyor type transport systems, algorithms for optimal control of the transport system flow parameters are required. Using neural network is the contemporary and perspective approach to synthesize such algorithms. The data sets required for training the neural network must be generated. This data sets must be generated in accordance with an experimental data. If this condition is not met, the use of these data sets for training a neural network is inefficient. In this paper, experimental data of the material input flow is analyzed. It is revealed that the hypothesis put forward about the normal law of distribution of the material input flow value does not correspond to the experimental data.

The existing recommendations that the correlation functions can be approximately represented in exponential form are confirmed. Additional studies are required of works that present experimental data on the magnitude of the input material flow for modern transport conveyors. The analysis and generalization of the obtained results will make it possible to form a countable number of relevant data sets which can be used to train neural network.

VIII. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] O. Pihnastyi and V. Khodusov, "Hydrodynamic Model of Transport System," *East European Journal of Physics*, vol. 1, pp. 121–136, 2020, doi: 10.26565/2312-4334-2020-1-11

- [2] Examples of use, <https://conveyorbeltguide.com/examples-of-use.html> (accessed Nov. 15, 2022)
- [3] "SIMINE – mining for the future," Mining industry | Market-specific Solutions | Siemens Global, <https://new.siemens.com/global/en/markets/mining-industry.html> (accessed Nov. 15, 2022)
- [4] O. Pihnastyi and V. Khodusov, "Model of a composite magistral conveyor line," *2018 IEEE International Conference on System analysis & Intelligent computing (SAIC)*, 2018, pp. 68-72
- [5] O. Pihnastyi, "Control of the belt speed at unbalanced loading of the conveyor," *Scientific bulletin of National Mining University*, 2019, no. 6, pp. 122-129, doi: 10.29202/nvngu/2019-6/18
- [6] I. Halepoto and M. Uqaili, "Design and Implementation of Intelligent Energy Efficient Conveyor System Model Based on Variable Speed Drive Control and Physical Modeling," *International Journal of Control and Automation*, 2016, no. 9 (6), pp. 379–388, doi: 10.14257/ijca.2016.9.6.36
- [7] H. Lauhoff, "Speed Control on Belt Conveyors—Does it Really Save Energy?" *Bulk Solids Handling*, 2005, no. 25(6), pp. 368–377
- [8] O. Pihnastyi, G. Kozhevnikov and V. Khodusov, "Conveyor Model with Input and Output Accumulating Bunker," *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 2020, pp. 253–258, doi: 10.1109/DESSERT50317.2020.9124996
- [9] P. Bardzinski, P. Walker and W. Kawalec, "Simulation of random tagged ore flow through the bunker in a belt conveying system," *International Journal of Simulation Modelling*, 2018, no. 4, pp. 597–608, doi: 10.2507/IJSIMM17(4)445
- [10] O. Pihnastyi, G. Kozhevnikov and M. Glavchev, "Energy Management Strategy in the Synthesis of an Algorithm for Multi-step Conveyor Belt Speed Control," *Lecture Notes in Mechanical Engineering*, 2022, pp. 33–42, doi: 10.1007/978-3-030-91327-4_4
- [11] R. Granell, C. J. Axon and D. Wallom, "Predicting winning and losing businesses when changing electricity tariffs," *Appl. Energy*, 2014, no. 133, pp. 298–307, doi: 10.1016/j.apenergy.2014.07.098
- [12] O. M. Pihnastyi, O. V. Ivanovska and M. O. Sobol, "Synthesis of the algorithm for the flow parameters optimal control of the reversible conveyor," *Radio Electronics, Computer Science, Control*, 2022, no. 2, pp. 210-221, doi: 10.15588/1607-3274-2022-2-20
- [13] A. Kirjanów, "The possibility for adopting an artificial neural network model in the diagnostics of conveyor belt splices," *Interdisciplinary issues in mining and geology*, vol.6, pp.1-11, 2016
- [14] D. Więcek, A. Burduk and I. Kuric, "The use of ANN in improving efficiency and ensuring the stability of the copper ore mining process," *Acta Montanistica Slovaca*, vol. 24, pp. 1-14, 2019
- [15] Xi Pingyuan and Song Yandong, "Application Research on BP Neural Network PID Control of the Belt Conveyor," *JDIM*, vol. 9(6), pp. 266-270, 2011
- [16] O. Pihnastyi and V. Khodusov, "Calculation of the parameters of the composite conveyor line with a constant speed of movement of subjects of labour," *Scientific bulletin of National Mining University*, vol. 4 (166), pp. 138-146, 2018
- [17] W. Kawalec and R. Król, "Generating of electric energy by a declined overburden conveyor in a continuous surface mine," *Energies*, vol. 14, pp. 1–13, 2021, doi: 10.3390/en14134030
- [18] V. Prokuda, "Research and assessment of cargo flows on the main conveyor transport “PSP Pavlogradskaya Mine, DTEK Pavlogradugol,” *Mining electromechanics*, vol. 88, pp. 107–111, 2012
- [19] V. Zaika, Yu. Razumny and V. Prokuda, "Regulated drives influence on coal flow and energy efficiency of mine conveyor transport system," *Naukovyi Visnyk Natsional. Hirnych. Univer.*, vol. 3, pp. 82–88, 2015
- [20] V. Kondrakhin, N. Stadnik and P. Belitsky, "Statistical Analysis of Mine Belt Conveyor Operating Parameters," *Naukovi pratsi DonNTU (Seriya elektromekhanichna)*, vol. 2(26), pp. 140–150, 2013

Use of State Pattern with The Ability to Dynamic Add New States for Software Part of Automatic Cleaning-In-Place Stations for Fermentation Tanks

Pavlo Katin, Oleksandr Pokhylenko

*Department of radio engineering systems
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine*

Abstract—This article has a brief summary of a research for partially or fully automated cleaning-in-place systems for washing largest fermentation tanks. For washing such tanks, the cleaning-in-place station of fermentation tanks (CIPSFT) are widely used. However, these systems have significant drawbacks. First is the need of human operator control of washing results, purity of solvents and water. Second is the need of program modernization for significant modifications. To solve this problem, a mathematical model in the form of a finite state machine and a State Pattern in object oriented paradigm for software CIPSFT controller were developed and described in this article. The proposed solution has the ability to quickly modify the program or change the program directly during the operation of the CIPSFT.

Keywords—*cleaning-in-place station of fermentation tanks, stages of the technological process of washing, generalized state diagram of the system, finite state machine with the ability to add new states.*

I. INTRODUCTION

Ensuring sanitary and hygienic requirements in the production process is one of the important tasks to be solved in the field of automation of food production. This applies to the production processes of beverages made in the fermentation process. For example, during the production of beer, contamination from fermentation residues remains in the working tanks. To remove them before and until now, manual cleaning is used. This method is advisable to use for small volumes of products. However, in the production of large volumes, manual cleaning of tanks is time-consuming and dangerous for personnel. According to [1-4], we list the volumes of the largest containers for beer production, for example: the CCT-32000C tank contains from 32000 to 38400 liters of beer, the CCT-SHP3-6000DE tank can hold from 6000 to 7200 liters of beer. There are also many other examples of the large volumes of tanks.

For washing such tanks, cleaning-in-place systems, which are partially or fully automated, for washing fermentation tanks are widely used. The cleaning-in-place station of fermentation tanks (CIPSFT) is designed for washing the inner surface of production equipment without the need for disassembly using hot alkali and acid. Unlike other known technologies, CIPSFT allows to reduce time and improve the quality of washing.

II. RELATED WORK

The work of the CIPSFT is carried out with the minimum participation of human operator. It can be used in production not only for cleaning fermentation tanks, but also for other objects connected to the system. In some cases, CIPSFT is a component of an automatic (automated) system of the full cycle of beer production in large volumes and can even use artificial intelligence methods [5]. We will give examples of the volume of tanks of the system of non-dismantling

washing of fermentation tanks: the volume of solution tanks up to 4000 liters, the volume of concentrate tanks up to 1000 liters, the volume of the heat exchanger up to 100 liters.

The most complex modern CIPSFT are controlled by specialized industrial controllers. These controllers must be programmed and serviced by the operator. In some cases, it is necessary to quickly modify the program or change the program directly during the operation of the CIPSFT. This is especially true during production modernization, when new fermentation tanks are added or taken out of usage.

The related work to improving the architecture of software, which is implemented on the basis of the state pattern, were considered in the next works [6-9].

The results for controller's software with state pattern was in [7, 9]. In these works, the theoretical part and practical of the state pattern issues and the hardware implementation of theoretical provisions are partially disclosed. But in [7, 9] the results don't have the ability to quickly modify the program or change the program directly during the operation of controller software.

An important results of state patterns for embedded systems with the possibility of dynamic creation of states was in [6]. This is working solution of problem to add new states during the operation of controller software.

Summarizing the above, it can be argued that the State pattern is widely used in programming, including software of system automation. To further outline the conducted and prospective studies in this article, we created general solution of this problem by developing and describing a mathematical model in the form of a finite state machine and a State Pattern in object oriented paradigm for software of CIPSFT controller, for ability to quickly modify the program or change the program directly during the operation of the CIPSFT or any system.

III. METHODOLOGY

Let's describe the essence of the problem in more detail. Nowadays there are 2 types of CIPSFT that are installed at factories. The division into types is carried out according to two main principles of using water and solutions.

The first type is CIPSFT, in which detergents and water for washing are used several times. Most often, such a CIPSFT contains 3-4 containers with water, an alkaline solution, a detergent, an acid solution and a drain tank. Such CIPSFT saves money, as it repeatedly uses both water and solutions. Energy conservation is also ensured. Control of cleaning results, purity of solvents and water is carried out by the operator in a semi-automated mode.

The second type of CIPSFT uses a new solution for each subsequent cleaning. In this case, production lines are cleaned with the same quality, and the sink does not accumulate dirt from previous procedures. However, costs increase because more clean water, solutions and energy are needed to prepare it.

Of course, in the conditions of the energy crisis and the lack of clean water, the first type of CIPSFT is relevant. However, these systems have significant drawbacks.

Firstly, it consists in the fact that the control of washing results, purity of solvents and water is carried out by the operator in a semi-automated mode. This can lead to deterioration of sanitary and hygienic characteristics in the production process. One of the reasons for this state of affairs is the complexity of the industrial controller software for CIPSFT and the difficulty of debugging the system as a whole.

Secondly, during the modernization of production, when new tanks are added, the program needs to be modified. This can especially cause complications when tanks of a new design are added and their automatic control requires a significant modification of the software part of the CIPSFT controller. In order to partially solve this problem, the article developed a mathematical model in the form of a finite state machine and a State Pattern with the possibility of dynamic adding new states.

1. Generalized description of the technological process of washing

To reveal the essence of the solution, it is necessary to describe the generalized work cycle of the CIPSFT and the mathematical model of it. In general, the system includes the following elements:

- washing station (CIPSFT) with tanks for solution, concentrates and clean water;
- cleaning facility (fermentation tanks);
- infrastructure elements of the cleaning system, which includes CIPSFT, heating system, pumps, valves, etc.

The analysis of existing solutions on the example of the scheme given in [2] made it possible to develop a generalized scheme of an automatic system based on CIPSFT (Fig. 1). This scheme does not take into account the heating subsystem in order to improve the investigation of the issue related to the State pattern, which does not reduce the overall progress of the investigation.

A station of this type usually has 3 working containers (Fig. 1): a container for circulating water (3), a container for an alkaline solution (4), and a container for an acid solution (5). In addition, CIPSFT has 2 tanks of concentrated solutions (Fig. 1): tank of concentrated alkaline solution (2) and tank of concentrated acid solution (1).

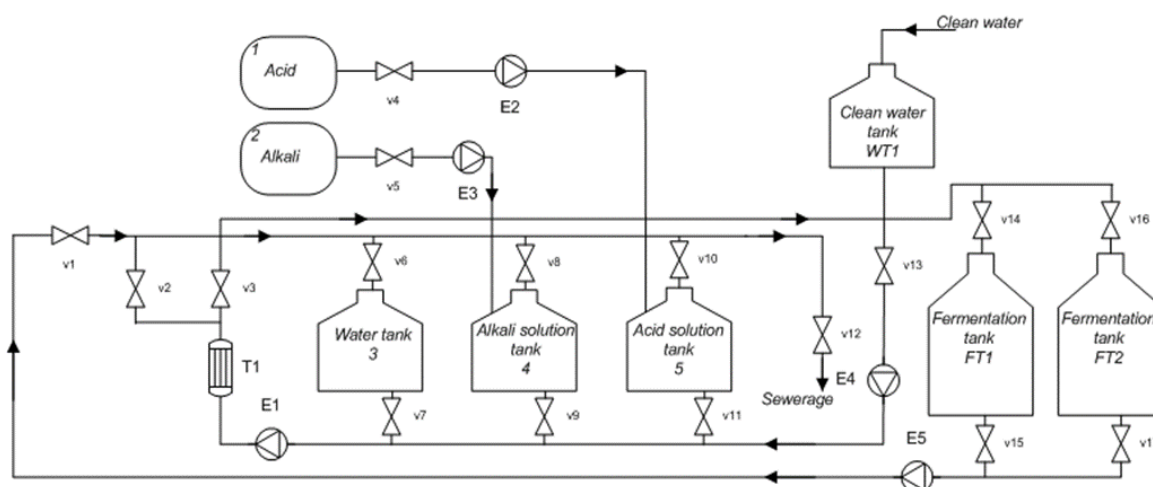


Fig. 1. Generalized scheme of the system based on CIPSFT.

The washing station is equipped with a heating system (T1, Fig. 1). One of the options for implementing heating is a tubular heat exchanger (TO), which heats solutions to a certain temperature. The tanks, heat exchanger and fermentation tanks are interconnected by a pipeline containing a system of sensors and valves for process control. Usually, the valves are controlled by the CIPSFT industrial controller. During fermentation, contaminations accumulate on the inner surface, and a viscous emulsion remains in the cone of the tank. Rotary washing heads are used for washing. A solution is supplied to the head under pressure, and the head makes rotational movements, which ensures that the solution is sprayed over the entire inner surface of the fermentation tank.

At the first stage of development, it is necessary to briefly describe the automatic system of the CIPSFT. As an example, we will describe the stages of the technological process of washing fermentation tanks.

2. Stage of washing FT with circulating water

At this stage, the content of WT under pressure is supplied to the cleaning object in the fermentation tank FT1 or FT2 to the washing head. Washing of the object is carried out by spraying circulating water under excessive pressure created by the pump in the circulation process. The pressure is provided by the supply pump E1 (Fig. 1).

At the first stage, the mixture goes through the following path: circulating water tank WT, valve V7, pump E1, heat exchanger T1, valves V3 and V14, FT1 tank, valve V15, pump E5, valves V1 and V6, circulating water tank WT. Thus, we have a closed system in which water circulates. The first stage is completed after the time specified by the controller. Process automation can be applied to improve the process. In this version, the results are checked using devices and sensors and the washing is finished.

3. Stage of washing FT with an alkaline solution

At the second stage, the content of alkaline solution tank AST is supplied under pressure to the cleaning object on the washing head. The object is washed by spraying an alkaline solution in the circulation process. The pressure is provided in the same way with the help of pump E1 (Fig. 1). The mixture goes through the following path: alkaline solution tank (AST4), valve V9, pump E1, heat exchanger T1, valves V3 and V14, tank FT1, V15, pump E5, valves V1 and V8.

4. The step of washing FT with an acid solution

At the third stage, the content of acid solution tank AST is supplied under pressure to the cleaning object on the washing head. The mixture goes through the following path: acid solution tank (AST5), valve V11, pump E1, heat exchanger T1, valves V3 and V14, tank FT1, V15, pump E5, valve V10.

If a lack or low quality of working solutions is detected, they are merged and a fresh solution is prepared.

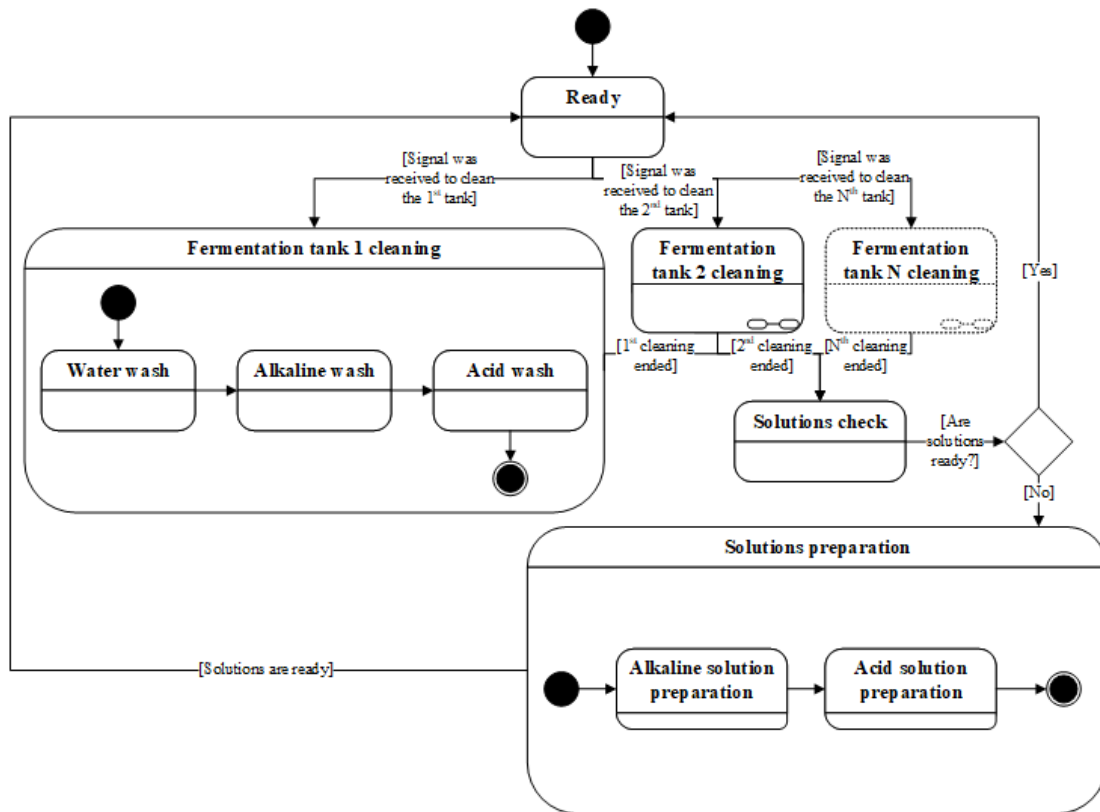


Fig. 2. Generalized state diagram of the system based on CIPSFT.

IV. RESULTS

A state diagram, which is shown in Fig. 2, was created to set the research task and clarify the problem.

Let's introduce certain restrictions for constructing the state diagram. Let's assume that there are no failures and malfunctions in the process of washing and operation of the system based on CIPSFT. So, we will assume that the system is deterministic. Let's divide the technological process of washing the fermentation tanks of the CIPSFT operation into separate stages, which is confirmed by the practice of using the CIPSFT in semi-automatic mode under the guidance of the operator. Such stages include: the stage of washing the fermentation tank with circulating water, the stage of washing the fermentation tank with an acid solution, the stage of washing with an alkaline solution, the technological process of preparing solutions.

At the same time, the technological process of preparing solutions includes two stages: the process of preparing an acid solution and the process of preparing an alkaline solution. The third limitation is that we will exclude the process of temperature regulation from the technological cycle of the system based on CIPSFT.

Since there can be several working tanks in the production, the washing system assumes that the washing takes place only for one tank at a certain time. Therefore, the operation of the CIPSFT system can be formalized in the form of a finite state machine, the number of states of which corresponds to the number of working tanks and is shown in Fig. 2.

The states of washing certain tanks contain nested states that formalize the technological stages of tanks washing, which are generally defined as three (Fig. 2). The state that formalizes the stage of preparation of the working solution also contains two internal states, namely: the state of preparation of the alkaline solution and the state of preparation of the acid solution (Fig. 2).

During production modernization, new fermentation tanks can be added. This is shown by a dotted line in Fig. 2. In this case, it is necessary to change the CIPSFT controller program. However, the proposed solution makes it possible to add new tank (new state) dynamically without changing the source code.

1. Variant of Software Implementation of a Finite State Machine with The Ability to Add New States

Considering the results of the work [6] and the state diagram from Fig. 2, it is possible to claim that the software part of automatic cleaning-in-place stations for fermentation tanks can be implemented using the State pattern with the ability to dynamically add new states. According to [6], the mathematical model of such a state machine is given by the following tuple:

$$M_3 = (Q_3, q, \Sigma_I, \Sigma_O, h) \quad (1)$$

with state space $Q_3 = Q_1 \cup Q_2 = \{q \mid q \in Q_1 \text{ or } q \in Q_2\}$, where Q_1 is initial state space and Q_2 is space of added states, initial state $q_0 \in Q_1$, finite input and output alphabets Σ_I, Σ_O , and transition relation $h \subseteq Q_3 \times \Sigma_I \times \Sigma_O \times Q_3$.

According to the formal description of the model (1) it is possible to match the elements of the solution proposed in [6] with a simplified class diagram which is shown in Fig. 3. This solution was practically implemented on microcontroller Cortex-M on C++ programming language.

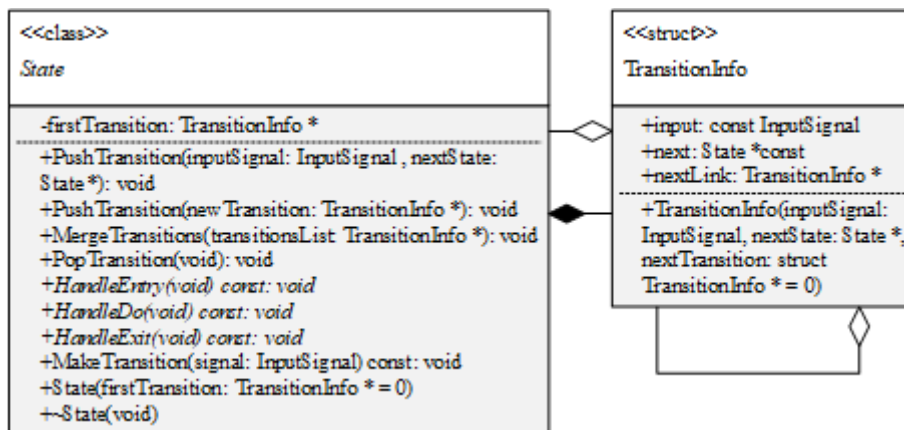


Fig. 3. Class diagram of the software implementation of the finite state machine with the ability to add new states.

Because the solution uses dynamic addition of states and transitions, the AbstractState class has methods for working with transitions: PushTransition, MergeTransitions and PopTransition that are used to add one transition, add several transitions and delete the last transition respectively. Using these methods leads to the change of transition relation h . To add a transition, the instance of the TransitionInfo structure must be created. This structure contains information about the input signal. This signal belongs to input alphabet Σ_I and leads to transition to next state from the set of states Q the pointer of which is stored in this structure. In the example under consideration, such signals can be received from a timer, a human operator or sensors when cleaning is completed.

Using the abstract State class, the new states may be added by creating the new subclasses of it. In this case the creation of the new instance of the class is equivalent to adding of a new element to Q_3 state space (1). The creation of the first instance of the class is equivalent to formation of initial state q_0 .

To perform useful operations for the complex embedded system the program implementation of corresponding virtual method is added to the subclasses of the abstract class State. Such useful operations are as follows: obtaining commands for executive mechanisms, voltage supply to general purpose input-output ports, passing signals to serial interfaces of microcontroller, reception of signals from sensors etc. Formally, they belong to output alphabet Σ_o . In the context of the current task, this allows to open and close valves, start and stop pumps and control the heating system.

2. Object-Oriented Programming Implementation

In [6] and in Fig. 3 the object-oriented programming (OOP) implementation is used. The speed of transitions between states and performing useful action was demonstrated in [7]. According to [7], the execution is the fastest in OOP implementation, so it demonstrates the feasibility of using OOP implementation for embedded systems. Thus, the proposed solution should be implemented in the OOP paradigm. If we decompose the state machine into basic states, then the instances of classes for the following tasks should be created:

- ready state – class with one instance, idle (doesn't perform actions, only waits for signal);
- water wash state – one instance per tank;
- alkaline wash state – one instance per tank;
- acid wash state – one instance per tank;
- solutions check state – class with one instance;
- alkaline solution preparation state – class with one instance;
- acid solution preparation state – class with one instance.

From this list, it is possible to see that to add a new tank, 3 instances of classes corresponding to the composite state should be created, and also 4 transitions should be added:

- from ready state to new water wash state;
- from new water wash state to new alkaline wash state;
- from new alkaline wash state to new acid wash state;
- from new acid wash state to solutions check state.

Thus, if the classes with the necessary logic have already been created, then adding a new fermentation tank is accompanied by only creating instances of the existing classes and the TransitionInfo structure, which can be done right during the system operation without stopping it.

V. DISCUSSION

The task of this work was to prepare a slight increase in the software architecture of CIPSFT controller to quickly modify the program or change the program directly during the operation of the CIPSFT. This problem was partially solved by developing a mathematical model in the form of a finite state machine and a State Pattern in object oriented paradigm.

At the same time there are wide distribute solutions of procedural programming paradigm. This is good result for embedded systems and industrial controllers. Therefore, these results should not be considered as a final decision.

VI. FUTURE RESEARCH

Future studies can address to processes of developing and testing of real software prototype for CIPSFT controller. Also the analysis of other standard patterns to create a solid software architecture of CIPSFT controller is needed.

VII. CONCLUSION

As a research results we created universal class diagram for the software implementation of the finite state machine with the ability to add new states. It is designed to create a software for cleaning-in-place station of fermentation tanks. This result has three advantages. The first is development using standard State pattern for software architecture of CIPSFT controller. The second is development of programs, algorithms and structures based on mathematical analysis. The third is ability to add new states in software architecture and quickly modify the program or change the program directly during the operation of the CIPSFT.

During the modernization of CIPSFT system, when new elements are added in this system, the program of system controller does not change at compilation level. This can especially cause complications when tanks or a new design are added in CIPSFT system. Their software does not require a significant modification. In this we partially solve this problem by developing a mathematical model in the form of a finite state machine and a State Pattern with the possibility of dynamic adding new states.

VIII. ACKNOWLEDGMENT

The authors gratefully acknowledge the professor, doctor of technical sciences Telenyk Sergii, for help in the actual research.

IX. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] "CCT-C Cylindrically-conical tanks CLASSIC – CBS." Czech Brewery System. <https://eshop.czechminibreweries.com/product-category/bpt/cct-ccf/classic/> (accessed Oct. 28, 2022).
- [2] "OFV Opened fermentation vats / Uncovered beer fermentors - Beer tanks." Beer Tanks. <http://www.beertanks.eu/offer/tanks/open-fermentation-vats/> (accessed Oct. 28, 2022).
- [3] A. Y. Tamime, Ed. *Cleaning-in-Place: Dairy, Food and Beverage Operations*, 3rd ed. John Wiley & Sons, 2009.
- [4] S. Salo, A. Friis and G. Wirtanen, "Improving the cleaning of tanks," in *Handbook of Hygiene Control in the Food Industry*. Woodhead Publishing, 2005, pp. 497-506.
- [5] S. Vassileva and S. Mileva, "Ai-Based Software Tools for Beer Brewing Monitoring and Control," *Biotechnology & Biotechnological Equipment*, vol. 24 no. 3, pp. 1936–1939, 2010, doi: 10.2478/v10133-010-0060-0.
- [6] O. A. Pokhylenko and P. Y. Katin, "State Pattern for Embedded Systems with The Possibility of Dynamic Creation of States," (in Ukrainian), *Technical Sciences and Technologies*, no. 1, pp. 118-127, 2021. [Online]. Available: http://nbuv.gov.ua/UJRN/tnt_2021_1_16.
- [7] P. Y. Katin and O. A. Pokhylenko, "Typical "State" Software Patterns for Creating Cortex-M Microcontroller System Software Infrastructure in Real Time Embedded Systems," (in Ukrainian), *Elektron. Model*, vol. 43, no. 2, pp. 51-67, 2021. [Online]. Available: <https://www.emodel.org.ua/en/archive/2021/43-2/43-2-5>.
- [8] V. O. Chmelov, P. Y. Katin and V. M. Shemaev, "Development of typical "State" software patterns for Cortex-M microcontrollers in real time," *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 9-105, pp. 29-38, 2020, doi: 10.15587/1729-4061.2020.205377.
- [9] P. Y. Katin, "Development of variant of software architecture implementation for low-power general purpose microcontrollers by finite state machines," *EUREKA: Physics and Engineering*, vol. 3, pp. 49–55, 2017, doi: 10.21303/2461-4262.2017.00361.

Behaviour of SPP waves frequency spectrum for dielectric/metal/dielectric structures

Petro Kostrobij, Vitalii Polovyi, Iryna Ryzha

*Department of Applied Mathematics
Lviv Polytechnic National University
Lviv, Ukraine*

Abstract—*in this paper, the authors studied the influence of different phenomena and materials properties on the frequency spectrum of the surface plasmon-polariton (SPP) waves behaviour in “dielectric/metal/dielectric” structures. Among them are spatial dispersion, quantization of energy levels, electron-electron interaction (Coulomb correlations) thickness of a metal layer. Established that for atomically thin metal films mentioned effects affect the frequency spectrum significantly. Mentioned that taking into account quantum-sized effects reveals a specific oscillatory picture of the spectrum that repeats one obtained for chemical potential. Proposed model yields results that have a quite good agreement with an experimental data.*

Keywords—*surface plasmon-polaritons, frequency spectrum, atomically thin metal films, quantum-sized effects, Coloumb correlations.*

I. INTRODUCTION AND RELATED WORK

The basic application of surface plasmon-polariton waves in so-called plasmon structures (i.e., structures in which propagate SPP) is the developing on their basis of compact highly sensitive optical waveguides [1,2], plasmon biosensors [3-5] designed to study the presence of impurities, the effect of drugs in a real time, etc. which work is based on the surface plasmons resonance effect [6] and laser stethoscopy devices [7].

Also, since plasmons are characterized by a high localization and intensity increase at the nanoscale, the control of the light at the nanoscale is a promising field. In addition, it should be noted that another important application of plasmonics is the use as a conductive material of a planar graphene [8-12], which is characterized by a high activity of charge carriers [10,13] and, therefore, leads to less energy losses [11,12,15].

Another promising application of plasmonics that will be of most interest to us in this paper is the use of metal-dielectric heterogeneous structures to excite plasmon waves, such as dielectric/metal, metal/dielectric or dielectric/metal/dielectric [1,15-23]. Usage of different materials for metals or dielectrics makes it possible to obtain plasmon waves with different integral characteristics, such as frequency or length of propagation.

Given the variety of application, various properties of SPP waves and existence of necessary tools for research, both mathematical and experimental, this topic also attracts the attention of many experimental scientists [22-27].

In this paper we focused our attention on a dielectric/metal/dielectric structures, when a metal is an atomically thin film (ATMF) with a thickness up to 100 nm.

II. PROBLEM FORMULATION

A structure we simulated propagation of SPP waves in consists of two dielectrics with sandwiched ATMP between. The coordinate for the structured is chosen so the XOY plane divides ATMP on two even parts. SPP wave will propagate along X axis. According to this we have the following mathematical boundaries for the areas:

$$\Omega = \begin{cases} \Omega_1, & z < -l/2, \\ \Omega_2, & -l/2 < z < l/2, \\ \Omega_3, & z > l/2. \end{cases} \quad (1)$$

Dielectric permittivity of insulators $\varepsilon_1(\omega)$ and $\varepsilon_3(\omega)$ are taken in the high-frequency approximation, thus, they are functions of the time variable/frequency. On the contrary, the dielectric permittivity of a metal depends on both time variable and spatial coordinates t, \vec{r} .

$$\varepsilon(\vec{r}, \vec{r}', t - t'). \quad (2)$$

Here we assume that external charges (stimuli) ρ in the area of ATMF are absent, so the Maxwell system of equations are such

$$\begin{aligned} \text{div} \vec{D} &= 0, \text{div} \vec{B} = 0, \\ \text{rot} \vec{H} &= \frac{\partial \vec{D}}{\partial t}, \text{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \end{aligned} \quad (3)$$

Here \vec{H} and \vec{E} are vectors of magnetic and electric strength and \vec{D} and \vec{B} are vectors of electric and magnetic flux density.

As we mentioned before in our model, we considered a spatial dispersion in the area of ATMF. In this case a connection between \vec{D} and \vec{E} vectors are non-local as process we investigate is non-stationary by nature [28]

$$\vec{D}(\vec{r}, t) = \int_{\Omega} d\vec{r}' \int_t dt' \varepsilon(\vec{r}, \vec{r}', t - t') \vec{E}(\vec{r}', t'). \quad (4)$$

Also, we considered a transverse magnetic (TM) polarization of vectors \vec{H} and \vec{E} vectors as for this type of polarization SSPs exists. Then we have the following form of these vectors:

$$\begin{aligned} \vec{H} &= (0, H_y, 0), \vec{E} = (E_x, 0, E_z) \\ \vec{H} &= H(z) e^{ik_x x} e^{-i\omega t}. \end{aligned} \quad (5)$$

After applying a Fourier transform to equations (3) and (4) and considering (5) we obtain a system of wave equations

$$\begin{aligned} \frac{\partial^2 H_y}{\partial z^2} + (k_0^2 \varepsilon_1(\omega) - k_x^2) H_y &= 0, \quad z < -\frac{l}{2}, \\ \frac{\partial^2 H_y}{\partial z^2} + (k_0^2 \varepsilon_2(z, \omega) - k_x^2) H_y &= 0, \quad -\frac{l}{2} < z < \frac{l}{2}, \\ \frac{\partial^2 H_y}{\partial z^2} + (k_0^2 \varepsilon_3(\omega) - k_x^2) H_y &= 0, \quad z > \frac{l}{2} \end{aligned} \quad (6)$$

The solution the equations for areas Ω_1 and Ω_3 is straightforward and simple. As for the equation in the ATMF Ω_2 we have an equation with a non-constant coefficient $k_0^2 \varepsilon_2(z, \omega) - k_x^2$ and, therefore, cannot be solved analytically. Thus, we needed to make some additional assumption about $\varepsilon_2(z, \omega)$ to proceed with a problem solution.

III. SOLUTION OF THE PROBLEM

Based on results obtained in [29] for a metal layer, we considered a model of an electron gas in a symmetric rectangular potential well of an infinite depth. The dielectric function based on the diagonal component of the dielectric permittivity tensor has the following form for this model:

$$\varepsilon_2(\omega, z) = 1 - \left(\frac{\omega_p^2}{2\pi n_e \omega^2} \sum_{i=1}^{n_{max}} (k_F^2 - \alpha_i^2) |\varphi_i(z)|^2 \right). \quad (7)$$

Here n_e – electron concentration in a metal, $\omega_p = \sqrt{4\pi n_e e^2 / m_e}$, k_F – the Fermi wave vector [30], α – quantum numbers, n_{max} – a number of bound states (energy levels).

The presence of surfaces is described by potential

$$U(z) = \begin{cases} \infty & \text{if } z \leq 0, z \geq l, \\ 0 & \text{if } 0 < z < l, \end{cases} \quad (8)$$

where l – width of a potential well. Function

$$\psi_n(z) = \sqrt{\frac{2}{S}} e^{i(\vec{q} \cdot \vec{r})} \varphi_n(z) \quad (9)$$

is a wave function of an electron in metal, $\vec{q} = (k_x, k_y)$, $\vec{r} = (r_{||}, z)$, $\varphi_n(z)$ – the solution of the Schrödinger equation with a Dirichlet boundary conditions.

$$\frac{\hbar^2}{2m} \frac{d^2 \varphi_n(z)}{dz^2} + U(z) \varphi_n(z) = W \varphi_n(z), \quad (10)$$

$$\lim_{z \rightarrow \pm\infty} \varphi_n(z) = 0. \quad (11)$$

W – is a full energy.

Equation (10), (11) with the potential (8) yield the expression for the wave function [29]:

$$\varphi_n(z) = \begin{cases} \sqrt{\frac{2}{l}} \sin(\alpha z) & \text{if } 0 < z < l, \\ 0 & \text{if } z \leq 0, z \geq l. \end{cases} \quad (12)$$

Quantum numbers α and maximum number of bound states n_{max} are given by relations:

$$\alpha_n = \frac{\pi n}{l}, \quad n_{max} = \left[\frac{l k_F}{\pi} \right], \quad k_F = \frac{\sqrt{2m\mu}}{\hbar}. \quad (13)$$

where $[\cdot]$ is a ceiling function and k_F is a magnitude of the Fermi wave vector, μ is a chemical potential [30].

It is known that for satisfying the conditions of electroneutrality the width of the potential well due to the presence of an exponential "tail" of the electron density in dielectrics near the surface of ATMF (so called quantum spill-out effect) does not coincide with the boundaries of ATMF and depends on the penetration depth of electrons into the dielectric [29]. As shown in [30]

$$l = l_{slab} + 2d, \quad (14)$$

$$d = \frac{3\pi}{8k_F} + \frac{\pi^2}{8k_F^2 l}. \quad (15)$$

is a function of k_F and a geometric thickness of a ATMF l_{slab} . Combining together (14) and (15) for l we have [30]:

$$l(k_F) = \frac{l}{2} + \frac{3\pi}{8k_F} + \frac{\sqrt{16k_F^2 l^2 + 24\pi k_F l + 25\pi^2}}{8k_F}. \quad (16)$$

It should be mentioned that in a proposed model we consider Coulomb correlations but only through their influence on the chemical potential μ and, consequently, the number of quantization levels n_{max} in the dielectric function of ATMF and neglecting their influence on other parts of the system.

Considering all the above we have an expression for a dielectric function $\varepsilon_2(z, z', \omega)$ in ATMF

$$\begin{aligned}\varepsilon_2(z, z', \omega) &= \varepsilon_2(z, \omega) \delta(z - z') = \\ &= \left(1 - \frac{\omega_p^2}{\pi n_e \omega^2} \sum_{n=1}^{n_{max}} \Sigma (k_F^2 - \alpha_n^2) |\varphi_n(z)|^2 \right) \delta(z - z').\end{aligned}\quad (17)$$

In [31] we showed that the dielectric permittivity function (17) and a wave function of electron (9) significantly depend on coordinates only at the dielectric-metal and metal-dielectric interfaces. Means that (17) significantly varies from constant only at boundaries of the potential well. So, we can make the following assumption about (17)

$$\varepsilon_2(z, z', \omega) \approx \varepsilon_2(l, \omega), \quad (18)$$

$$\begin{aligned}\varepsilon_2(l, \omega) &= \frac{1}{l} \int_0^l \varepsilon_2(z, \omega) dz = \\ &= 1 - \frac{\omega_p^2}{2\pi n_e \omega^2} \sum_{n=1}^{n_{max}} (k_F^2 - \alpha_n^2) |\bar{\varphi}_n(z)|^2,\end{aligned}\quad (19)$$

$$|\bar{\varphi}_n(z)|^2 = \frac{1}{l} \int_0^l |\varphi_n(z)| dz. \quad (20)$$

In this case, we limited ourselves to considering the first equation of the system

$$\frac{\partial^2 H_y(z)}{\partial z^2} + (k_0^2 \varepsilon_2(l, \omega) - k_x^2) H_y(z) = 0. \quad (21)$$

When the wave equation has such form, the dispersion relation can be easily found and has the following form [31]

$$e^{-4k_1 \frac{l}{2}} = \frac{k_1 - k_2 \frac{k_3}{k_2} \frac{k_2}{k_1}}{\frac{\varepsilon_1}{k_1 + k_2} \frac{\varepsilon_3}{\varepsilon_2} \frac{\varepsilon_2}{\varepsilon_3} \frac{k_2}{k_1}} \quad (22)$$

$$k_i = \sqrt{k_x^2 - k_0^2 \varepsilon_i}. \quad (23)$$

IV. NUMERICAL ANALYSIS

The equation (22) is a transcendental algebraic equation that should be solved numerically. The Newton's method [32] was applied for roots finding. The same method was applied for finding a chemical potential μ from the electroneutrality condition. As we already mentioned here, we are interested in finding solutions (in this case SPPs frequency spectrum) dependence on a metal thickness l . So, logically, we decided to make an initial loop over the thickness with a step equal to 0.01 a.u. In this loop we firstly found a chemical potential and, therefore, a magnitude of the Fermi wave vector k_F and energy levels n_{max} . After that we were able to find a potential well width from equation (16) also using the Newton's method. After establishing all the necessary values on the previous step, we iterate over wave vector k_x absolute value with a step equal to 0.04 a.u. In this loop we solve the equation (22) for a frequency ω and eliminate physically useless solutions (e.g., where $\text{Re}(\omega) < 0$ or where $\text{Re}(\omega)$ or $\text{Im}(\omega)$ are uncharacteristically high).

V. RESULTS

Study of the obtained model were conducted for such structures "SiO2/Ag/Si", "Vacuum/Ag/Si", "Vacuum/Ag/Al2O3" and, characteristics of the metal are taken from [29] and of the dielectrics from [33]. The results of simulation are shown in Fig. 2-4 and Tab. 1.

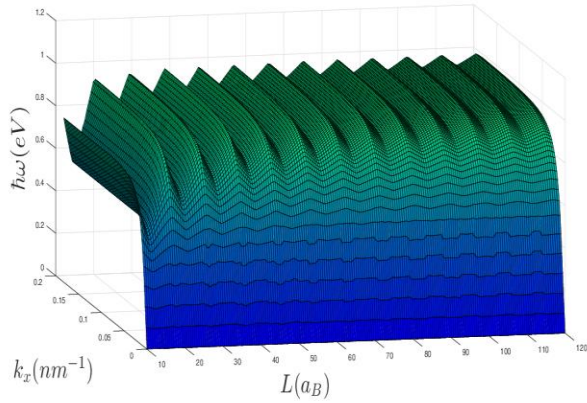


Fig. 1. Dependence of plasmon spectrum on metal thickness for “Vacuum/Ag/Si” structure (with the Coulomb correlation).

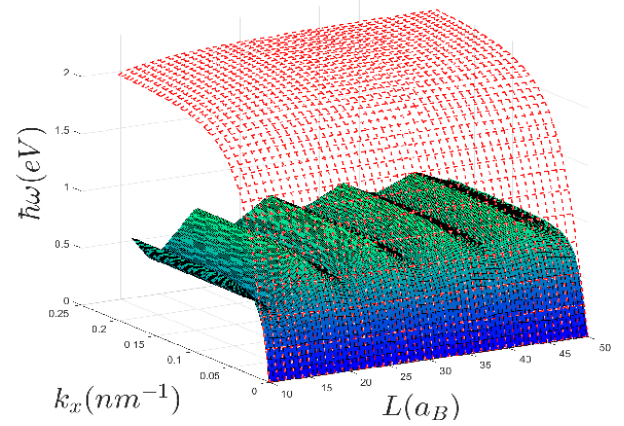


Fig. 2. Frequency spectrum for “SiO₂/Ag/Si”. Without the Coulomb correlations - red surface, with the Coulomb correlations - blue surface.

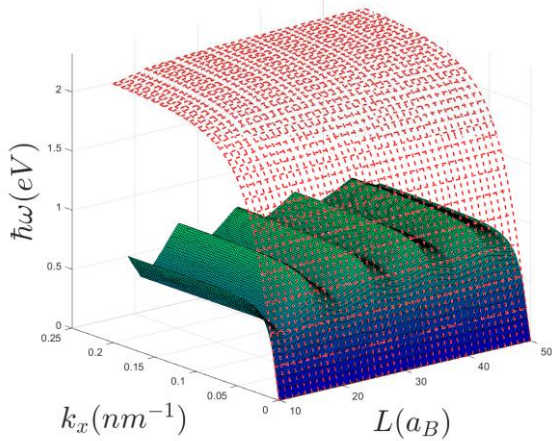


Fig. 3. Frequency spectrum for “Vacuum/Ag/Si”. Without the Coulomb correlations - red surface, with the Coulomb correlations - blue surface.

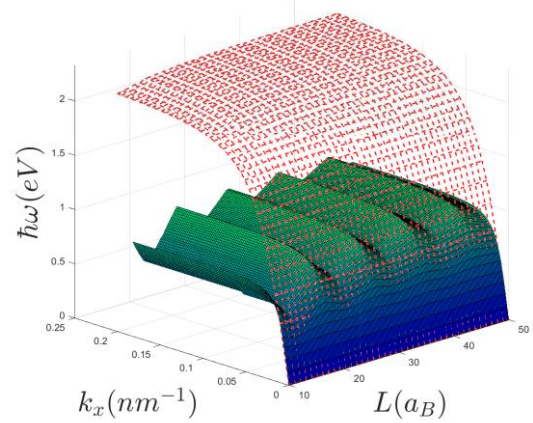


Fig. 5. Frequency spectrum for “Vacuum/Ag/Al₃O₂”. Without the Coulomb correlations - red surface, with the Coulomb correlations - blue surface.

TABLE I. COMPARISON OF THE MODEL SIMULATION WITH AN EXPERIMENTAL DATA

k_x (nm ⁻¹)	A (eV)	B (eV)	E (eV)
0.0278	0.696	1.288	0.62
0.0487	0.889	1.789	0.8

VI. CONCLUSION

The results show that even a rather "rough" consideration of the Coulomb correlations, namely, their influence on the chemical potential leads to significant changes in the frequency spectrum of SPPs compared to the classical approach and correct considering of the condition of electroneutrality for a non-interacting system of electrons. This is explained by the fact that in the case of an interacting system, the influence of quantum size effects increases significantly [29]. The oscillation pattern of the spectrum with an increase in the film thickness becomes less noticeable Fig.2, which coincides with the corresponding behavior of a chemical potential [30], the decay speed of the oscillation peaks also largely depends on the Wigner-Seitz radius. It is worth noting the dependence of the spectrum on the dielectrics surrounding ATMF which can be seen by comparing the data given in Fig.3,4,5.

It is important to note that the inclusion of Coulomb correlations also leads to a significant improvement in the agreement with the experimental data Tab.2, and this proves the need to take these correlations into account when modeling the processes of propagation of SPPs waves in ATMF.

VII. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] S.A. Maier, Springer – Verlag, *Plasmonics: Fundamentals and Application*, 2007.
 - [2] M. Ono, T. Hideaki et al., *NTT Technical Review*, **16**, 7, 14-19, 2018.
 - [3] J.R. Mejía-Salazar, N.O. Oliveira, American Chemical - Chemical Reviews, Society, *Plasmonic Biosensing*, **118**, 2018.
 - [4] J. Liu, M. Jalali et al., *Analyst - The Royal Society of Chemistry*, **145** 2, 364-384, 2020.
 - [5] H. Šípová, S. Zhang, *Anal. Chem.*, **82**, 24, 10110–10115, 2010.
 - [6] J. Nico, J. Marcel, E. Fischer, *Surface Plasmon Resonance*, **627**, ISBN: 978-1-60761-669-6, 2010.
 - [7] C. Chen, D. Mohr et al., American Chemical Society - Nano Letter, **12**, 12, 7601-7608, 2018.
 - [8] M. Jablan, H. Buljan, M. Soljačić, *Phys. Rev. B* **80**, 245435, 2009.
 - [9] D. C. Pedrelli, B. S. C. Alexandre, *EPL* **126** 27001, 2019.
 - [10] L. Xiaoguang, Q. Teng, et al., *Materials Science and Engineering: R: Reports*, **74**, 11, 351-376, 2013.
 - [11] A. Principi, M. Carrega et al, *Phys. Rev. B* **90**, 165408, 2014.
 - [12] L. H. Zeng, C. Zhang et al., *Sci Rep* **5**, 8443, 2015.
 - [13] K.I. Bolotin, K.J. Sikes et al., *Solid State Communications*, **146**, 9–10, 351-355, 2008.
 - [14] S. V. Boriskina, T. A. Cooper et al., *Advances in Optics and Photonics*, **9**, 4, 775-827, 2017.
 - [15] E. E. Narimanov, *Phys. Rev. A* **99**, 023827, 2019.
 - [16] D. Ziemkiewicz, K. Słowik, and S. Zielińska-Raczyńska, *Opt. Lett.* **43**, 490-493, 2018.
 - [17] Qi Zhang, Chaohua Tan et al., *Chin. Opt. Lett.* **13**, 082401, 2015.
 - [18] S. Ali Hassani Gangaraj and F. Monticone, *Optica* **6**, 1158-1165, 2019.
 - [19] T. Tomohiro et al., *J. Phys.: Condens. Matter* **31** 305001, 2019.
 - [20] A. Marini, D. V. Skryabin, and B. Malomed, *Opt. Express* **19**, 6616-6622, 2011.
 - [21] V. G. Achanta, *Reviews in Physics*, **5**, 100041, 2020.
 - [22] M. Abd El-Fattah, Zakaria & Mkhitaryan, Vahagn & Brede, et al., *Plasmonics in Atomically-Thin Crystalline Silver Films*, 2019.
 - [23] A. R. Echarri, J. D. Cox, F. J. Garcia de Abajo, *Quantum Effects in the Acoustic Plasmons of Atomically-Thin Heterostructures*, [arXiv:1901.07098v3], 2019.
 - [24] L. Zheng, U. Zywietz, A. Evlyukhin, *Sensors (Basel)* **19** (21):4633, 2019.
 - [25] L. Grave de Peralta, D. Domínguez, *Optics Communications*, **286**, 151-155, 2013.
 - [26] L. Lin, *Manipulation of Near Field Propagation and Far Field Radiation of Surface Plasmon Polariton*, Springer Theses, 2017.
 - [27] Gao, Xi and Cui, Tie Jun, *Nanotechnology Reviews*, **4**, 3, 239-258, 2015.
 - [28] J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, 2007.
- 2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

- [29] P. Kostrobij, B. Markovych, Philosophical Magazine Letters, *Effect of Coulomb interaction on chemical potential of metal film*, 2018.
- [30] N. Ashcroft, N. Mermin, Cornell University, Harcourt, *Solid State Physics*, 1976.
- [31] P. Kostrobij, V. Polovyi, MMC, **6**, 2, 297-303, Influence of the Thickness of a Metal Nanofilm on the Spectrum of Surface Plasmons, 2019.
- [32] C. T. Kelley, *Solving nonlinear equations with Newton's method*, Fundamentals of algorithms, SIAM, Philadelphia, PA, ISBN 0-89871-546-6, 2003.
- [33] V. Pogosov, A. Babich, and P. Vakula, Phys. Solid State **55**, 2120, 2013.

Risks Assessment and Approaches to Creative of the Reliable Software Modules for IoT Devices

Vadym Malinovskyi, Leonid Kupershtein, Vitaliy Lukichov

Information Protection Department
Vinnitsia National Technical University
Vinnitsia city, Ukraine

Abstract — The some especial aspects and of software functional stability in the IoT devices and their risk factors and algorithmic stability were considered in this paper. The main risk factors and their probabilities, including special informational influences in IoT are considered in these paper. Also was considered a basic approaches of the reliable and fault tolerance algorithms and software modules.

Keywords — Internet of Things (IoT), Data stability, Informational influences, Software, Cybersecurity, Realiability analysis, Software quality, Software modules.

I. INTRODUCTION

Most users and industrial IoT and BYoD devices are more and more covers all spheres of humans live, from personal everyday use to professionals use and utilization in the industrial area – such of the Industrial Internet of Things (IIoT) and Commercial Internet of Things (CIoT) [1]. Now, in difficult times and hybrid war and world crisis the question of cyber security of xIoT devices is very acute and actual, what we can see at the big number of cyber attacks on IoT sector and computer and network sector in general. The latest trends of cyber threats are shown, that the individual specials types of cyber threats is a very dangerous and make made a software and hardware faults, which cause to economical and time losses [1]. A wide implementation of xIoT technologies of remote monitoring and remote control in industrial area, the consequences of cyber threats can be very significant and commensurate with the weapons use. That's making very actual to developing and implementation a various approaches and methods of software cyber defence and reliability increasing. Also it's allow organizing highly efficient, secure and comfortable automated management, monitoring and data processing of IoT information systems.

The goal of research – is a detailing a research of data stability in IoT software and provide and describes of more decision risks assess and approaches to achieve of maximum level of IoT software stability.

II. RELATED WORK

A related works provides a description of the main approaches to protection and processing and risk compensation in IoT data systems section is essential to most research authors articles [1] – [2], [4] – [7]. In this works were described a few key parameters, such as probability of failure-free operation, reliability function, and reliable parameter. But in some cases, given the difference in the intensity of cyber threats and non-compatibility of individual events of cyber threads – the some indexes of keys parameters $k(x)_i$ (such as probability of failure-free operation [1]) may be a very small compared to more significant cyber threats and unreliable factors. That it allows neglecting an individual, especially not influencing additive (or multiplicative) components of probabilities of failure-free operation [1]. That's, allows us to assume, that the some key parameters, such as probability was that the IoT or BYoD object will fail during time t characterizes the opposite property – unreliability [1] and is expressed, if taking to account only is a maximum probabilities, as: $q_{sum}(t) = \sum_{i=0}^N p_i(t) 1 - \frac{p_i \max(t) \gg p_i(t)}{p_i \max(t)}$. Obviously, $q_{sum}(t)$ $q(t)$ can be considered as a maximum failure distribution function factor (or sum of maximum, most important

probability), its derivative: $f_{i \max(t)} = -dpi^{\max(t)}/dt$. It's the density of the maximum distribution functions of uptime or, the density of maximum failures factors. Experimentally, as a results of the simulation in MathCad and MathLab environments, were received the values of maximum failure distribution function factor in the range: 0.7325 ... 0.9341 (high risk conditions) – in taking into account the conditions only of most important factors of cyberthreats and their influences; 0.1521 ... 0.4243 (low risk conditions) – in taking into account the conditions only smallest factors of cyberthreats. The average range of values of maximum failure distribution function factor is in range 0.4423 ... 0.6827. The average range of values of the probabilities of failure-free operation, which was receive by simulation is in range 0.52 ... 0.74 , that which is ensured in conditions with most important factors of cyber threats and uses reliable approaches in software.

Modern users and industrial IoT devices has a significant problems – it's a complex cybersecurity and low reliability for it's functionality, which slows down their implementation in the critical and industrial spheres. With the growing popularity of smart devices IoT services, the intensity of cyber threats and reducing of summery reliability is inversely increasing.

The trends of 2022 are indicate that the main problems of functionality in modern IoT are:

- Complex cybersecurity of IoT devices and combined with them devices;
- Different reliability risks of IoT devices and their modules;
- Software and hardware core components reliability and low fault free interferences each for each, at of operating functionality witch implements of the IoT platphorm total functionality and also functionality of complex data infrastructures with this IoT devices.

II. METHODOLOGY

The methodology in a research paper is the section involves the evaluation of the most significant ones. Also risks assessments are include only the main influence factors of most intensive cyber threats.

Risk assessment when working with information in the Internet of Things is carried out in accordance with the criterion of a comprehensive assessment using the likelihood of the occurrence of a threat:

$$P(\lambda_i) \max = \sum_{i=1}^m p_{\max i} \cdot k_i \cdot k_m = P_i(\lambda) \cdot K_{CMS}, \quad (1)$$

where p_i – unit probabilities of the occurrence of a cyber threat for each of the main informational factors of threats; $p_{\max i}$ – maximum unit probabilities of the compatible occurrence of a cyber threat for each of the main informational factors of threats. k_i – correction coefficients for each threat factor; K_S – complex correction coefficients for all threat factors; k_m – coefficient of risk level for each cyber thread or unreliabalance in each of units; K_{CMS} – complex risk factors coefficient .

It's a takes into account only the most important and most influence factors.

It's should be noted that for a comprehensive assessment of the cyber threats it will be fair $0 < P(\lambda_i) \max < 1$, and the higher the set of factors with the corresponding probabilities r_i , the greater the total probability will approach 1.

Accordingly, the conditions (1) of stable work (and implements a basic conditions, as shown in paper [1]) with the absence of risks are ensured by the evaluation of the indicator (stability coefficient):

$$R' \xrightarrow{t_i \rightarrow t_{i \min}} \frac{1}{P_{\max}(\lambda_i)} \rightarrow 1. \quad (2)$$

The higher the indicator R' ($R' \in 0 \dots 1$) – the better the conditions for the stability of the software in IoT information systems [1].

The most important data in IoT (financial, economical or important technical data, such as shown at the fig.1) and their sub parameters must be protected first and by main priority.



Figure 1. – Example Illustration : Protection the most important and critical sensitive data parameters in IoT devices

Fig. 1 is represents the main field of potential risk of most important data modification and importance of it's security and protections in data procession IoT software and also is each of it's unit.

The principles of reliable functionality for operating units in IoT software are provides by some separate approaches in software development and implementations methods in IoT. First of all, must be provided a high level of cyber security and hardware reliability by different ways and tools. Second, must be released reliable and reservation data algorithms (data graphs) in critical software, compatible with high reliable and stable hardware support (reliable realizable hardware architecture and components).

Third, must be granted a traditional and famous data model of strong separate parameters and separate access levels and rights of user data with complex security methods are widely used to protect access to resources. A number of modern attacks and modern hacking software can use many methods to explore and hack user data and software modules (or their units) and get access to it's in IoT devices. That needs to gain unauthorized access to internal software, which can then be modified or downloaded. Especially in current times, the hierarchy and evolution of methods and software systems of both a technical nature and social engineering for obtaining closed data and information for the purpose and further use of them for fraudulent and cybercriminal purposes and committing crimes of an economic nature has significantly increased.

To best protect data and increase its stability and functionality it is important to understand what makes this data and data software units reliable, how it content can be identified by a unique identifier and reproduce and verify at another point (mechanism of checksums/Hash functions) with complex protection mechanisms. A special feature is also the authorization data and ensuring the reliable implementation of complex authorization models (such as Triple AAA model, provided by Cisco Inc.). In particular it create the reliable accounts and reduce a parts of risks, when accessing the software part and as one of the components of the security policy information systems as a whole. Creating a reliable checksum (hash function) mechanisms on different stages of computing processes for each software modules, and also reliable strong passwords are an integral part of the security measures of modern data systems. Trusted identifiers and data must meet the requirements listed in order of importance:

1. Have unchanging unique content and functionality;
2. The value of identifiers does not change as at the point of creation information data, as well as at intermediate points and to the final points of their reception/processing;
3. Work fault tolerance algorithms should not take a lot of resources and not have extremes of work functions (peaks during work in tracking mode);
4. Methods of conversion and comparison of keys of data identifiers should not be complex and take up a lot of computing resources.
5. Uses an mechanisms of automated controls of check checksum and hash function of software modules by all its lifecycle and working process. Also it's may take places for the working files and functional data of this software modules.
6. Identifiers must not be readily accessible to functional testing modules and must be protected from other leaks and access. Hash functions and checksums should be relatively complex as well data and generated on the basis of the received data, may contain numbers, symbols and their mixture in the upper and lower registers.

The popularity of using trusted data and the checksum/Hash function mechanism to verify data integrity is due to the fact that such verification is simple to implement in digital hardware, easy to analyze, and well suited for detecting common errors and vulnerabilities, weak cryptoresistance caused by the presence of noise in data channels, and information systems.

Given such a large number of potentially possible cyber threats and information risks for IoT and mobile personal devices, it is necessary to use comprehensive IoT approaches and mechanisms at all levels. It is also relevant to develop new progressive approaches and world-leading practices, such as demarcation of networks, IoT segments, ZeroTrust area, data protection systems for IoT. Basic model was shown in paper [1]. The use of a comprehensive method of checking and neutralizing cyber threats is also relevant. In general, the structural mechanism and complex approach to data protection in the Internet of Things and its component should include the parallel use of information protection mechanisms:

$$F_{Actual}(IoT\ Security) \rightarrow F_{Max}(t_b, x_i \in n; y_i \in m; z_i \in k; t_i \rightarrow t_{imin}, p_i \rightarrow p_{max}) + \Delta F(F(t_b, x_i \in n; y_i \in m; z_i \in k; t_b, p_i) + F_{ZeroTrustZonePolicies}(t_b, x_i \in n; y_i \in m; z_i \in k; t_i \rightarrow t_{imin}, p_i \rightarrow p_{max})) \quad (3)$$

where, $F_{Actual}(IoT\ DataSecurity)$, $F_{Max}(IoT\ DataSecurity)$ – actual, max and additional (from minimal risk factors components) designation of a complex conditional function of maximum IoT information protection and reliability indexes with a minimum number of threats in IoT systems; $F_{ZeroTrustZonePolicies}(t_b, x_i \in n; y_i \in m; z_i \in k; t_i \rightarrow t_{imin})$ – the use security functions by providing of access rights delimitation policies and information security policies based on the concept of zero trust in IoT zones; t_b – conditional time intervals; $x_i \in n; y_i \in m; z_i \in k$ – corresponding information parameters and their belonging to sets; $t_i \rightarrow t_{imin}$ – criteria for performing functions in the shortest possible time.

In general, it is possible to achieve the maximum level of protection (minimal risks factors) in the Internet of Things devices only with the use of an integrated complex approaches, with consideration only max influence factors in data flow model, the use of the above-mentioned individual complex components and integrated approach for information protection function in IoT, under conditions:

$$F_{Actual}(IoT\ DataSecurity) \rightarrow F_{Max}(IoT\ DataSecurity) + F_{\Delta}(IoT\ DataSecurity) \quad (4)$$

It is extremely difficult to ensure full functional security and secure data transmission and processing for personal IoT with mobile personal devices of users as part of it, taking into account the different functional orientation and the use of individual multi-structured components in the complex and multi-component information system of modern IoT, as well as taking into account the specifics of the use of publicly available Internet channels – as one of the main sources of cyber threats.

Ensuring stability and reliability of functionality, the concept of data integrity, availability and confidentiality (CPA) in modern IoT is one of the priority tasks on future. New models and methods should be based on a complex combination of functionality, data virtualization technologies, the use of modern IDS/IPS with mixed additional functionality. Also, in order to increase the level of security, additional conditions for checking and controlling third-party information flows with reliable improved encryption with offset and in combination with computing parallelism should be created process with demarcation of access rights at different levels of computing and virtual computing environments (shells) for different processes.

Approaches to Creative of Reliable and Fault Tolerance Algorithms must provides by strong reliable parallel processing or alternative logical condition paths processing if software irrational or failure condition are occurs. Fault tolerance end point (point of entry) in cycle must include a alternative way of processing in software algorithm path (algorithm path) or provide a start a spare software resources. Also may be involved a special additional software modules to provide a alternative ways to similar functions of data processing. That's provides a high level of functional stability of software. If the hardware fault is occur, and hardware architecture having a strong serial data processing model, that are not prevent to general software failure. That, for strong reliable stable and reliable data processing model in IoT software must be granted a parallel or reservation hardware architecture or fault tolerance or recovery methods [3].

III. RESULTS

Since informational threats in microprocessor tracts in data systems, including microprocessor devices of the Internet of Things (IoT) are quite often complex and have a complex nature and stages of implementation, then solutions aimed at protecting the computing process and the algorithms of the microprograms of the control microcontroller must also has a comprehensive approaches. This approaches of increasing the data security and reliability must realization by transmission and processing data flows in MC microprograms with more reserve and stability principles in IoT software.

To ensure the closure of potentially dangerous critical places of the IoT controller architecture, individual and complex approaches are used to organize the necessary state of security:

1. Control of the integrity and reliability of the memory content in software (including a strong control a sensitive and potentially malicious content), which is provided by checking and correcting errors of the Error Correction Code and checking parity. It also provides additional protection against attacks aimed at preventing code bugs from infecting systems;

2. Control of external and internal data flows and key-parameters of software platform in IoT. For example, a temperature sensor continuously measures the temperature of the environment surrounding the microcontroller, which also may be threatened;

3. Approaches, which involving the use of isolation and control of the integrity software modules by Hesh-functions and Checksum mechanisms (MD4, MD5 or SHA 256/SHA 512, CRC8, CRC16, CRC32 or others Hesh-functions algorithms). It's can also be used by cyclic redundancy code engineering or implementation of the fault tolerance algorithms with reserve alternative branches;

4. Hardware-based approaches that use a cyclic redundancy check calculate, i.e. a checksum is calculated that detects errors in data transmission or storage. Not only does this ensure code integrity is checked, but it also means that the signature can be calculated at runtime;

5. Monitoring KPI parameters of software life indexes (health indexes) and resource monitoring is another method with a high degree of protection. To determine the cause of the reset and thereby ensure reset only through authenticated access to the 'Cyberheat indicate flags' status management system.

In the results of work were received the dependencies in probabilities of failure-free operation and parameter of failure distribution function factor in some especially conditions.

Experimentally, as a results of work were received the values of the density of the maximum distribution functions and probabilities of failure-free operation. Maximum failure distribution function factor are in the range: 0.7325 ... 0.9341 (high risk conditions) – in conditions with only of most important factors of cyberthreats and their influences and that range of values are in range in 0.1521 ... 0.4243 (low risk conditions) – in conditions with smallest factors of cyberthreats influence. The average range of

values is in range 0.4423 ...0.6827. Also was received the average range of values of the probabilities of failure-free operation, which was receive by simulation is in range 0.52 ... 0.74 , that which is ensured in conditions with most important factors of cyber threats and uses reliable approaches in software.

IV. CONCLUSION

In the conclusion, providing of the maximum stability in IoT software may be reached by providing maximum cyber security and reliability of each software modules component in program environment in IoT. New models, such reliable and fault tolerance algorithms may increase a summary stability of IoT software and data processing model. Also, may take place a providing another perspective methods of stability increasing of IoT software complexly with hardware stability architecture and hardware methods. The given results in paper can to take into account an some most important factors of cyber heats and failures, and summery assessments of risk in IoT more precision.

REFERENCES

- [1] Malinovskiy Vadym, Kupershtein Leonid, Lukichov Vitaliy “Cybersecurity and Data Stability Analysis of IoT Devices”2022 *International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkiv 2022, pp. 474-478, [online] Available: <https://easychair.org/smart-slide/slide/IRtj>
- [2] V.V.Sklyar, V.V.Yatskiv, N.G.Yatskiv,“Dependability and SecurityInternet of Things: Practicum”, *Ministry of Education and Science of Ukraine, National Aerospace UniversityKhAI, Ternopil National Economic University*, 2019.
- [3] K. Nelson, R. Davis, D. Lutz, and W. Smith, "Optical generation of tunable ultrasonic waves," *Journal of Applied Physics*, vol. 53, no. 2, Feb., pp. 1144-1149, 2002
- [4] C. Canella, J. Van Bulck, M. Schwarz, M. Lipp, B. von Berg, P. Ortner, F. Piessens, D. Evtushkin, and D. Gruss, “A systematic evaluation of transient execution attacks and defenses,” *arXiv preprint arXiv:1811.05441*, 2018, [online] Available:<https://doi.org/10.48550/arXiv.1811.05441>.
- [5] O. Voitovych, Y. Baryshev, E. Kolibabchuk and L. Kupershtein, “Investigation of simple Denial-of-Service attacks”, *2016 Third International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)*, 2016, pp. 145-148, [online] Available:<https://doi.org/10.1109/INFOCOM-MST.2016.7905362>.
- [6] O. Voitovych, L. Kupershtein, O. Shulyatitska and V. Malyushytsky, “The authentication method in wireless sensor network based on trust model”,*2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2017, pp. 993-997, [online] Available: <https://doi.org/10.1109/UKRCON.2017.810039>.
- [7] O. Voitovych, L. Kupershtein, V. Lukichov and I. Mikityuk, “Multilayer Access for Database Protection”*2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, 2018, pp. 474-478, [online] Available:<https://doi.org/10.1109/INFOCOMMST.2018.8632152>.

Математична модель інформаційного забезпечення екологічних служб

Михайло Крихівський, Любов Саманів

*Кафедра прикладного програмування та обчислення
Івано-Франківський національний технічний університет нафти і газу
Івано-Франківськ, Україна*

Анотація—Дослідження направлене на розвиток методів математичного моделювання інформаційного забезпечення, розроблення інструментів об'єктивізації служб екологічного контролю у контексті концепції інформаційного моделювання та прогнозування. Запропоновано математичні моделі для покращення інформаційного забезпечення діяльності в розрізі розумних змін існуючих ситуацій екологічних і соціальних систем. Як комплексний підхід запропоновано враховувати користь і ризики, що поєднує майбутні здобутки та можливі втрати від прийняття управлінського рішення. Дослідження базується на формалізації нечіткості, неточності, невизначеності стану та процесів екологічних і соціальних системах.

Ключові слова—*Математична Модель Інформаційного Забезпечення, Прийняття Рішення Екологічними Службами, Нечіткий Стан Екосистеми, Нечіткий Екологічний Процес.*

I. ВСТУП

Найефективнішим способом підвищення рівня екологічної безпеки є управління нею, тобто підвищення екологічної безпеки у встановленому масштабі та за встановлених умов. Проте зрозуміло, що для планування, контролю та сприяння покращенню екологічної безпеки (управління цим збільшенням), перш за все необхідно створити доцільне інформаційне забезпечення екологічних служб. Для цього має бути доступна повна та достовірна інформація про умови навколишнього середовища в кількісному вигляді, який найбільш зручно використовувати в сучасних системах прийняття рішень.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Інформаційне забезпечення екологічних служб використовується для отримання достовірної інформації про фактичний стан компонентів екосистеми, що є багаторівневою [1]. На рівні підприємства джерелами інформації є засоби екологічного контролю підприємства, система моніторингу навколишнього середовища, комплекс нормативно правових документів щодо лімітів забруднення: статистичні звіти, квартальні звіти; дозволи на водокористування; дозволи на землекористування; сертифікати та дозволи на виробничу діяльність; документи про допустимість шкідливих речовин у техногенних забрудненнях; екологічні паспорти промислових підприємств. Контроль за виконанням вимог для сталого розвитку та забезпечення ефективного виконання комплексних екологічних заходів вимагає від різних екологічних служб уточнення операційної взаємодії та ефективності діяльності, оброблення відповідної інформації [2]. Глибоке розуміння тем і методів екологічної діагностики дозволяє розробити основні характеристики різних етапів роботи для забезпечення сталого розвитку [3]. Формальний опис інформаційного забезпечення базується на єдиному принципі побудови специфікації умови забезпечення необхідних умов [4] для полегшення подальшої алгоритмізації.

III. МЕТОДИ

У роботі використані дослідження структури статистичної інформації, яку оприлюднює Державна екологічна інспекція України, та її інформаційне забезпечення. Проаналізовано
2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

можливість використання методів математичного моделювання для покращення умов прийняття управлінських рішень.

IV. РЕЗУЛЬТАТИ

Для інтелектуалізації розрахунків пропонується інформаційна аналітична система, основою якої є семантична мережа, що реалізована нечітким орієнтованим зваженим графом з розміченими вершинами та ребрами. Скінчений орієнтований зважений граф \tilde{G} для реалізації математичних операцій за допомогою комп'ютера описано матрицею суміжності \tilde{S} і вектором ваг D . Прийняття рішення полягає у знаходженні номера найкориснішої альтернативи. У формальному описі проблеми задані множини виборів A й альтернатив B . Потужність множини виборів $|A| = n$, а потужність множини альтернатив $|B| = m$. Корисність кожної альтернативи задана послідовністю дійсних чисел, з яких формується матриця. Розв'язком задачі прийняття рішення є альтернатива із B , якій відповідає номер рядка матриці, тобто натуральне число, яке менше або рівне $m \in \mathbb{N}$. З метою врахування ризиків знаходиться сумарне значення ризику альтернатив за допомогою формули: $ra_i = \sum_{j=1}^n r_{ij}$, $i = 1, 2, \dots, m$ де r_{ij} – ризики, які визначаються формулою $\|r_{ij}\| = \left\| \max_i k_{ij} - k_{ij} \right\|$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$. Норма ризику $\|ra_i\|$ ($i = 1, 2, \dots, m$) альтернативи є фактично її сумарним ризиком, розділеним на максимальне значення із сумарних ризиків всіх альтернатив $\|ra_i\| = ra_i / \max_i ra_i$.

V. ОБГОВОРЕННЯ

Деякі об'єкти (стани об'єктів) екосистеми або деякі відношення можуть бути нечіткими або не визначеними, тому граф \tilde{G} є нечітким (розпливчастим). Множини \tilde{V} – нечітка множина, елементи якої визначені парою, (v, μ_v) , де v – елемент множини \tilde{V} , μ_v – значення функції належності цього елемента до множини \tilde{V} . Рядки та стовпці матриці суміжності \tilde{S} визначені парою (v, μ_{vi}) , де i – номер рядка чи стовпця, а μ_{vi} – значення функції належності до множини \tilde{V} вершини з номером i . Значення функції μ_{vi} для всіх i обмежені відрізком $[0, 1]$. Значення елементів матриці \tilde{S} , які є значеннями функції належності μ_A , також обмежені відрізком $[0, 1]$, а значення елементів вектору D обмежені відрізком $[-1, 1]$.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальшим удосконаленням розвитку дослідження нам представляється використання динамічного декомпозиційного підходу до опису змін структури екосистеми та математичних моделей складної гри з динамічною (силовою) взаємодією та перенесенням інформації в складних ієрархічно організованих системах.

VII. ВИСНОВКИ

У результаті теоретичних досліджень розроблено математичну модель комплексу інформаційного забезпечення екологічних служб, що є інтелектуальною основою для прийняття ефективного управлінського рішення.

ЛІТЕРАТУРА

- [1] Ю.В. Буц, «Систематизація процесів пірогенної релаксії екогеосистем в умовах техногенного навантаження» в «ЕКОЛОГІЧНА БЕЗПЕКА» № 1/2018 (25), 2018, с. 7-12, doi: 10.30929/2073-5057.2018.1.7-12.
- [2] О. Melnyk, «The analysis of the eu regional approaches to solving the ecological safety problems» в «ЕКОЛОГІЧНА БЕЗПЕКА» № 1/2017 (23), 2017, с. 7-13. [Online]. Available: <https://www.kdu.edu.ua>.
- [3] Л.О. Семець, «Адміністративно-правові засади управління у сфері забезпечення екологічної безпеки держави: напрями удосконалення» в «ГОСПОДАРСЬКЕ ПРАВО ТА ПРОЦЕС» №4, 2020, с. 101-107, doi: 10.32842/2078-3736/2020.4.14.
- [4] В. О. Копанчук, «Екологічна безпека як складова національної безпеки України: сучасні концепції та підходи» в «Вісник НАДУ» № 2, 2020, с. 45-49, doi: 10.36030/2310-2837-2(97)-2020-45-49.

Application of Blockchain and the SimHash Algorithm to Detect Plagiarism

Roman Dorosh

*Department of Information Technology
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine*

Abstract—In this paper, a novel, unified approach for detecting plagiarism was proposed. It is built on the Ethereum blockchain environment and uses the SimHash algorithm for similarity detection, scalability and plagiarism detection in education field as main area of application. Through a network of social and smart contracts, Blockchain technology will function as an open technology that protects against unwanted access, but at the same time will save timestamp and current state of the analyzed documents. Also, the unified algorithm was compared to widely used TF-IDF algorithm on the same dataset and showed improvement by nearly 7%.

Keywords— *Similarity Detection; SimHash; Blockchain; Smart-contracts; Education; Natural Language Processing; Ethereum Virtual Machine*

I. INTRODUCTION

Blockchain also provides the main characteristics: autonomy, decentralization, fault-tolerance, distribution, openness, immutability to previous version of the network and transparency [1], which led to interest of researchers into this area and current state of research is poor. Blockchain as a foundational technology is useless without being built upon. Building apps that interact with data on distributed nodes rather than just appending to databases is made possible by smart contracts, which enable programmatic support for blockchain. Everybody may access the content and data of the public blockchain since smart contracts are executed there. However, unless another version is deployed to a public network of connected nodes, they are immutable.

This article also offers a high-level description of the system, a plagiarism detection algorithm with support for different languages using word embeddings, and a smart contract that was implemented in a test EVM (Ethereum Virtual Machine). Also included in this study are real-world results, a detailed description of the method used to compare techniques using blockchain and SimHash, comparisons to other approaches, and a future outlook for this initial effort.

II. RELATED WORK

Following the release of Satoshi Nakamoto's initial paper [2], which was primarily intended to demonstrate how to obtain consensus through the decentralization of servers, blockchain began to acquire prominence. Thus, "miners" that serve the blockchain on their own hardware or use cloud-based services are rewarded for sustaining the infrastructure. The PoW (Proof of Work) algorithm is a type of algorithm that creates transactions that cannot be modified without repeating the PoW.

Following the popularity of Bitcoin and the realization of the limitations of this method of creating blockchains, Ethereum [3], a second widely used system based on PoS (Proof of Stake), emerges. Ethereum is a unique environment that enables the development of decentralized applications utilizing smart-contracts as rule-based systems and an API. Developers now have the ability to create web applications and publish them on the Ethereum blockchain.

Natural language processing is a common method for extracting relevant data from documents and normalizing it. Then, the simple cosine similarity, a slightly modified version of the cosine similarity [4], or the TF-IDF algorithm may be used to determine the degree of similarity between the texts. Neural networks, in particular, and machine learning may be used in fields like software defined demodulation of weak radio signals [5], effective identification document recognition [7], and weak radio signal demodulation. Research [8] demonstrates the significance of using novel educational strategies, particularly for online engineering education.

A Blockchain-Based Non-Fungible Tokens strategy with various similarity metrics was suggested by Pungilă et al. [9]. They come to the conclusion that their method outperforms conventional similarity assessments in terms of speed and is applicable to real-world circumstances. However, there are other downsides, including lots of created NFT (which translates into high prices) and the inability to recognize minimally changed versions of the same data that are presented in multiple ways (using the SHA256 hash).

III. METHODOLOGY

Scalability and speed are the SimHash algorithm's main benefits. The original study [10] demonstrates its practical applicability for locating near-duplicate documents in huge file storage (multi-billion repository). SimHash was therefore selected as the platform for plagiarism detection based on its implementation areas and properties. According to this method, we determine the SimHash of each element in the set of documents $S = \{D_1, D_2, D_3, \dots, D_n\}$ by looking through the documents (D). Then, each entry's hash SH_d is compared to each document S individually based on the distance between them. This procedure will produce a number in the range $[0, 1]$, which may be converted to the relevant percentage. Additionally, the Sh set of hashes is saved on the Ethereum blockchain via a smart contract and cannot be removed by a centralized authority.

SimHash technique provides ability to arrange similar papers into groups and identify many instances of plagiarism with a single database query. The biggest drawback in this situation is that the algorithm cannot accurately display the locations where the cheater copied from the original document. To facilitate multi-language plagiarism detection, the text processing consists of the following steps: tokenization, stop-word removal, stemming, lemmatization, and transfer to word embeddings. In this research, I suggest combining the decentralized blockchain technology and SimHash, two techniques of plagiarism detection, into a single algorithm (Figure 1)

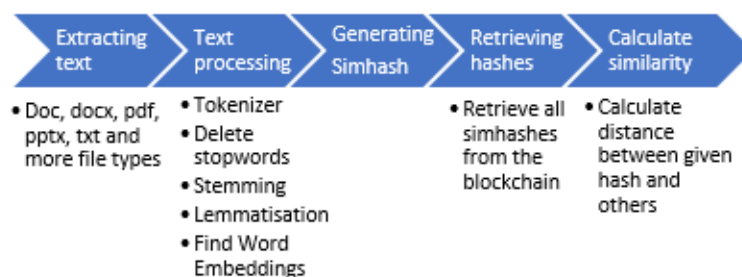


Figure 1 – Unified Algorithm

Finding plagiarism and calculate the similarity is not the easy task, especially when it comes to the multi-language comparison. For evaluation of the algorithm performance and accuracy I've implemented the next pseudo-algorithm (Figure 2)

Algorithm 2 Evaluation similarity pseudo algorithm

```

for each added Document do
  process Document
  generate SimHash for processed Document
  if generatedSimHash can be found on the Blockchain then
    added Document totally plagiarized ( $S \leftarrow 1$ )
  else
    calculate similarity  $S$ 
  end if
return 10 most common SimHashes and their  $S$  (similarity) score in range [0, 1]
end for

```

Figure 2 – Evaluation similarity pseudo algorithm

The outer loop makes request to the server with appropriate document for future analysis and processing. If the generated SimHash for given document is equal 1, it means that this document is totally plagiarized. If not, the proposed algorithm calculated similarity between current document and the all documents that available on the blockchain and return the 10 most “similar” with their scores. Also, the threshold for the document identified as plagiarized is equal to 25% and can be set programmatically.

IV. RESULTS

Precision and recall are the primary performance indicators for information retrieval, pattern recognition, and classification activities. While recall (sensitivity) is the percentage of relevant documents that were recovered, or alternatively, the question of how many relevant documents were obtained, precision (positive predictive value) is the fraction of positive detect plagiarism papers to all the retrieved documents. Other metrics and techniques exist for measuring the effectiveness of an information system, such as area under the ROC curve (AUC), which is also taken into account in this study. The dataset [11] for evaluation proposed in this paper consists of 143,000 news articles from 15 major publications but used only 10,000. The language of the articles is English, because multi-language support is not fully implemented. Also, I can formalize this task, as a classification with two classes (plagiarism - 0 and not plagiarism - 1).

I have performed testing on Processor 11th Gen Intel(R) Core(TM) i7-1185G7 @ 3.00GHz, 2995 Mhz, 4 Core(s), 8 Logical Processor(s) with 32 GB RAM and 512 GB SSD drive, OS is Windows 10. The Python library for visualizing metric results *sklearn* was used. So, the results of testing my algorithm is given below on the Figure 3 and Figure 4 (ROC AUC score).

	precision	recall	f1-score	support
0	0.90	0.90	0.90	4988
1	0.90	0.91	0.90	5012
accuracy			0.90	10000
macro avg	0.90	0.90	0.90	10000
weighted avg	0.90	0.90	0.90	10000

Figure 3 – Metrics of proposed algorithm

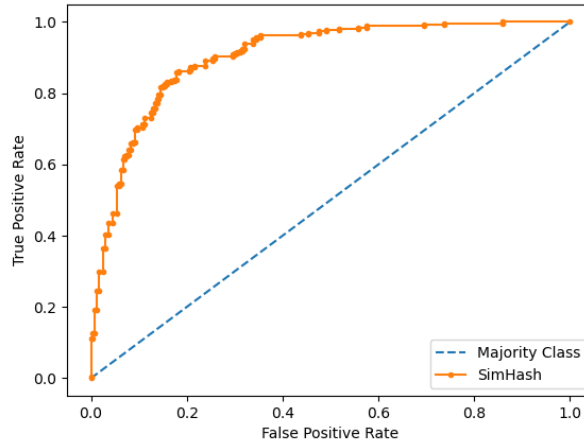


Figure 4 – ROC AUC score of proposed algorithm

V. DISCUSSION

In this work, I suggest a brand-new technique for detecting plagiarism that makes use of the Ethereum blockchain and SimHash. My algorithm has decent precision, recall, and f1-score based on the findings; however, it is not suitable for usage in production. Overall, it performs 7% better than the well-known algorithm TF-IDF.

Additionally, the system displays flawless results when an exact replica of copied content is submitted, but this is a very simple example. The algorithm detects some percentage of plagiarism when genuine and original work is contributed, but it does not provide false-positive results. The most difficult test instance is when a cheater plagiarizes original material and attempts to hide it by paraphrasing the entire line. As a result, the suggested approach outperforms the TF-IDF model while having the lowest success rate in detecting it among the three test instances. As a result, I must update my corpora with related and synonym words in order to detect paraphrasing and incorporate more intricate multilingual support.

VI. FUTURE RESEARCH

Since this algorithm was tested on the dataset presented only in English, detection between different languages is also important in further development. Also, the algorithm was compared only with the TF-IDF algorithm. Further research is possible for more constructive pre-processing and application of text models of machine learning.

VII. CONCLUSION

The proposed unified algorithm performs well on detecting exact or partial copies of the documents; additionally, the performance does not degrade with scalability and size of the documents; changing words to their synonyms or paraphrasing can also be detected; identification of multi-language plagiarism performs on the basic level; complex language analysis must be performed to achieve a more precise solution. Furthermore, the SimHash algorithm performs well on enormous numbers of documents (billions), therefore the pace of plagiarism detection is consistent between low and high document numbers.

REFERENCES

- [1] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli and M. H. Rehmani, "Applications of Blockchains in the Internet of Things: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1676-1717, Secondquarter 2019, doi: 10.1109/COMST.2018.2886932.

- [2] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", Aug 2022, [online] Available: <https://bitcoin.org/bitcoin.pdf>.
- [3] V. Buterin, "Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform", Aug 2022, [online] Available: https://ethereum.org/669c9e2e2027310b6b3cdce6e1c52962/Ethereum_Whitepaper_-_Buterin_2014.pdf.
- [4] Anzelmi, Daniele, Domenico Carlone, Fabio Rizzello, Robert Thomsen and Dil Muhammad Akbar Hussain. "Plagiarism Detection Based on SCAM Algorithm". In *Proceedings of the International MultiConference on Engineers and Computer Scientists 2011* (Vol. Volume I, pp. 272-277). Newswood Limited, International Association of Engineers, IAENG
- [5] M. Kozlenko, I. Lazarovych, V. Tkachuk and V. Vialkova, "Software Demodulation of Weak Radio Signals using Convolutional Neural Network," *2020 IEEE 7th International Conference on Energy Smart Systems (ESS)*, 2020, pp. 339-342, doi: 10.1109/ESS50319.2020.9160035.
- [6] M. Kozlenko and V. Vialkova, "Software Defined Demodulation of Multiple Frequency Shift Keying with Dense Neural Network for Weak Signal Communications," *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2020, pp. 590-595, doi: 10.1109/TCSET49122.2020.235501.
- [7] M. Kozlenko, V. Sendetskyi, O. Simkiv, N. Savchenko, A. Bosyi, "Identity Documents Recognition and Detection using Semantic Segmentation with Convolutional Neural Network (short paper)". *Cybersecurity Providing in Information and Telecommunication Systems 2021 (CPITS)*, 2021, pp. 234-242.
- [8] Dutchak, M., Kozlenko, M., Lazarovych, I., Lazarovych, N., Pikuliak, M., Savka, I. "Methods and Software Tools for Automated Synthesis of Adaptive Learning Trajectory in Intelligent Online Learning Management Systems". *Innovations in Smart Cities Applications Volume 4. SCA 2020. Lecture Notes in Networks and Systems*, vol 183. Springer, Cham. https://doi.org/10.1007/978-3-030-66840-2_16.
- [9] C. Pungilă, D. Galis, V. Negru, "A New High-Performance Approach to Approximate Pattern-Matching for Plagiarism Detection in Blockchain-Based Non-Fungible Tokens (NFTs)", arXiv:2205.14492 [cs.CR], May 2022.
- [10] G.S. Manku, A. Jain, A.D. Sarma, "Detecting Near-Duplicates for Web Crawling", research.google.com, <https://static.googleusercontent.com/media/research.google.com/en/pubs/archive/33026.pdf> (accessed Aug. 23, 2022).
- [11] "All the news dataset". Kaggle.com <https://www.kaggle.com/datasets/snapcrack/all-the-news> (accessed Aug. 27, 2022).

Аналіз та розпізнавання спаму в соціальних мережах та месенджерах

Василь Олійник, Андрій Подорожняк, Наталія Любченко

*Навчально-науковий інститут комп'ютерних наук та інформаційних технологій
Національний технічний університет "Харківський політехнічний інститут"
Харків, Україна*

Анотація—Існує велика кількість різних соціальних мереж і месенджерів, які в часи пандемії коронавірусу і російської війни в Україні займають велику частину всього нашого життя, особливо в роботі. Проблема боротьби зі спамом і спамерами є як ніколи актуальною, оскільки частка спаму в робочому текстовому потоці постійно залишається достатньо великою. У доповіді на прикладі відомих датасетів із застосуванням антиспам-боту у популярному месенджері Telegram запропоновано вирішення науково-прикладної задачі визначення спаму та спамерів у текстовому контексті месенджерів соціальних мереж з використанням оригінального комплексного мажоритарного алгоритму. Досягнуті значення F-міри точності тесту класифікації спаму для запропоновано комплексного мажоритарного алгоритму склали 0,999 на SMS Spam Collection Dataset та 0,949 на Spam Mails Dataset.

Ключові слова—Спам, Визначення спамерів, Соціальна мережа, Антиспам-бот, Комплексний мажоритарний алгоритм.

I. ВСТУП

Зазвичай прості чати в соціальних мережах та месенджерах не мають вбудованого алгоритму захисту від спаму, тільки електронні поштові скриньки мають таку функціональність. Це може бути причиною того, що відсоток спаму в таких чатах переважно достатньо великий. Нажаль невеликі компанії не мають можливості придбати спеціальний пакет месенджерів для ведення робочої бесіди, тому зазвичай вони використовують такі безкоштовні месенджери як наприклад Телеграм. Шкідливе посилання, вставлене в повідомлення та надіслане співробітникам компанії, може становити велику небезпеку для всієї компанії. Тому в сучасному світі постала проблема моніторингу вхідного текстового потоку в соціальних мережах і месенджерах. Також необхідно ідентифікувати та блокувати спамерів, що пришвидшує роботу алгоритмів і ускладнює життя спамерам.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Найнадійніший спосіб боротьби зі спамом – це не дати спамерам отримати вашу електронну адресу. Програмне забезпечення для автоматичного виявлення спаму називається фільтрами захисту від спаму. Фільтри захисту від спаму можуть застосовуватися кінцевими користувачами або на серверах. Це програмне забезпечення має два основні підходи [1].

1. Зміст повідомлення аналізується, на підставі чого робиться висновок, спам це чи ні [2]. Якщо повідомлення класифіковано як спам, його можна позначити або навіть видалити. Таке програмне забезпечення може працювати як на сервері, так і на клієнтському комп'ютері. За такого підходу спам не відфільтровується, але ви продовжуєте сплачувати повну вартість його отримання, оскільки антиспам програмне забезпечення все одно отримує кожне спам-повідомлення (марно використовуючи ваші комп'ютерні ресурси та витрачаючи гроші) і лише потім вирішує, показувати його чи ні.

2. Відправник класифікується як спамер, не дивлячись на текст повідомлення. Таке програмне забезпечення може працювати лише на сервері, який безпосередньо отримує повідомлення. За допомогою такого підходу можна зменшити витрати – гроші витрачаються лише на зв'язок із програмами розсилки спаму (тобто на відмову приймати повідомлення) і на зв'язок з іншими серверами (якщо такі є) для перевірки. Виграш, однак, не такий великий, як можна було б очікувати. Якщо одержувач відмовляється прийняти повідомлення, спамерська програма намагається обійти захист і відправити його іншим шляхом. Кожну таку спробу доводиться відбивати окремо, що збільшує навантаження на сервер.

У доповіді також розглядається декілька основних методів виявлення спамерів. Зазвичай наявні параметри виявлення спаму поділяють на дві групи: на лінгвістичній основі, та на основі поведінки користувачів соціальних мереж та меседжерів [3, 4]. Лінгвістичні методи спрямовані на виділення дискримінаційних лінгвістичних ознак, щоб відрізнити фальшивих користувачів від звичайних. Наприклад, ці методи ідентифікують спам-рецензії відповідно до лінгвістичної підказки, особливості стилю написання, синтаксичного шаблону, моделі теми на основі LDA, генеративної моделі Байєса, позитивного немаркованого навчання, моделі на основі фрейму та функцій на рівні документа.

Виявлення спаму на основі поведінки має на меті виявлення набору колективних зловмисних маніпуляцій онлайн-оглядами відповідно до ознак поведінки. Пропонований нами алгоритм пов'язаний із лінгвістичним виявленням спаму, оскільки ми визначаємо, чи є користувач спамером, на основі його повідомлень [5, 6].

III. МЕТОДИ

Було використано три базові найпопулярніші алгоритми розпізнавання спаму: наївний байєсівський класифікатор [7], метод опорних векторів [8], згорткову нейронну мережу [9, 10]. Пропонований метод на базі комплексного мажоритарного алгоритму полягає у тому, що після отримання повідомлення від користувача соціальної мережі чи месенджера (у нашому випадку від користувача Telegram), якщо користувач невідомий у нашій системі, то він додається до нашої бази даних (БД) разом із усіма користувачами програми. Після чого аналізуються повідомлення користувача із використанням перелічених вище алгоритмів. На основі отриманих даних обчислюється відсоток спаму в повідомленні та передається результат в аналізатор спаму. Аналізатор спаму вирішує, чи є користувач, який надіслав повідомлення, спамером, чи ні, на основі визначеного відсотка спаму в поточному повідомленні та двох останніх прогнозів [6]. Отже, щоб ідентифікувати користувача як спамера, аналізуються його три останні повідомлення, і якщо середній відсоток спаму в його повідомленнях перевищує вказане значення, ми користувач визнається як спамер і його ідентифікатор додається до БД зі спамерами .

IV. РЕЗУЛЬТАТИ

Через велику несиметричність відомих наборів даних спам повідомлень SMS Spam Collection Dataset [11] та Spam Mails Dataset [12] замість звичайної метрики точності для оцінки роботи вибраних алгоритмів, ми використовували оцінку міри F1. Також для чистоти тестування ми використовували два різних датасети.

Результати тестування алгоритмів на навчальних і тестових вибірках датасету SMS Spam Collection Dataset:

Algorithm	Training sample	Test sample
Bayes	0.988	0.982

SVM	0.998	0.989
CNN	0.990	0.985
Majority	1.000	0.999.

Результати тестування алгоритмів на навчальних і тестових зразках датасету Spam Mails Dataset:

Algorithm	Training sample	Test sample
Bayes	0.922	0.898
SVM	0.983	0.956
CNN	0.954	0.949
Majority	0.965	0.959.

V. ОБГОВОРЕННЯ

Уважно оглянувши результати можливо зрозуміти, що найбільшу точність дає саме запропонований комплексний мажоритарний алгоритм, тобто коли ми аналізуємо данні одночасно усіма запропонованими вище алгоритмів. Мінус цього алгоритму полягає в тому, що час реакції такого способу більше ніж у випадку використання лише одно з алгоритмів. Але через те, що ми спираємося не тільки на результати одного алгоритму та не тільки на одне повідомлення, ми зменшуємо помилку блокування звичайного користувача соціальної мережі чи месенджера як спамера. Це потрібно брати до уваги тому, що в залежності від того, як саме ми обрахуємо кінцевий результат спираючись на данні отримані від алгоритмів, результати дослідження можуть змінюватися. Наприклад, у різних випадках ми можемо віддавати перевагу тому чи іншому алгоритму в залежності від особливостей прикладної області.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В майбутньому планується проаналізувати більшу кількість алгоритмів та їх комбінацій, використати трохи інші методи обчислення кінцевого результату та ввести так звані коефіцієнти до кожного із застосованих алгоритмів. Нажаль на даний момент технічні можливості з проведення таких досліджень суттєво обмежені. Також планується проведення тестування на інших датасетах із інших прикладних областей, а також створення свій датасету, спираючись на текстовий контент одного конкретного телеграм чату.

VII. ВИСНОВКИ

У рамках доповіді на прикладі декількох датасетів з використанням антиспам-боту у популярному месенджері Telegram вирішено науково-прикладну задачу визначення спаму та спамерів у текстовому контексті месенджерів соціальних мереж.

Крім того, розглянуто актуальність виявлення спаму та можливі проблеми через втручання спаму у соціальні мережі та месенджери, проаналізовані основні методи розпізнавання спаму, а саме наївний байєсівський класифікатор, метод опорних векторів, згортова нейронна мережа, запропоновано власний комплексний мажоритарний алгоритм та створені їх програмні реалізації.

Також приведено основні методи виявлення спамерів, та запропонований власний оригінальний метод виявлення спамерів на базі аналізу його трьох останніх повідомлень.

Розроблено програму для фільтрації спаму та виявлення спамерів у месенджері Telegram, яка використовує три популярні алгоритми розпізнавання спаму (наївний байєсівський класифікатор, метод опорних векторів, згорткова нейронна мережа) та запропонований авторами комплексний мажоритарний алгоритм.

ЛІТЕРАТУРА

- [1] S. Chaudhry, S. Dhawan, and R. Tanwar, "Spam Detection in Social Network Using Machine Learning Approach," *Communications in Computer and Information Science*, vol. 1230, pp. 236-245, May 2020, doi: 10.1007/978-981-15-5830-6_20.
- [2] N. Liubchenko, A. Podorozhniak, and V. Oliinyk, "Research of antispam bot algorithms for social networks," *CEUR Workshop Proceedings*, vol. 2870, pp. 822-831, Apr. 2021. [Online]. Available: <http://ceur-ws.org/Vol-2870/paper61.pdf>.
- [3] D. Koggalahewa, Y. Xu, and E. Foo, "An unsupervised method for social network spammer detection based on user information interests," *J. of Big Data*, 9, 7, Jan. 2022, doi: 10.1186/s40537-021-00552-5.
- [4] Y. Meleshko, O. Drieiev, and H. Drieieva, "Method of identification bot profiles based on neural networks in recommendation systems," *Advanced Information Systems*, vol. 4, no. 2, pp. 24-28, Jun. 2020, doi: 10.20998/2522-9052.2020.2.05.
- [5] V. Oliinyk, N. Liubchenko and A. Podorozhniak, "Spam recognition and spammers detection," *Proceedings of the 9th international scientific and technical conference Problems of informatization*, Ukraine, 2021, vol. 1, p. 46, Nov. 18-19, 2021. [Online]. Available: http://repository.kpi.kharkov.ua/bitstream/KhPI-Press/54913/1/Conference_NTU_KhPI_2021_Problemy_informatyzatsii_Ch_1.pdf.
- [6] N. Liubchenko, A. Podorozhniak, and V. Oliinyk, "Research Application of the Spam Filtering and Spammer Detection Algorithms on Social Media," *CEUR Workshop Proceedings*, vol. 3171, pp. 116-126, May 2022. [Online]. Available: <http://ceur-ws.org/Vol-3171/paper13.pdf>.
- [7] S. Sugahara, and M. Ueno, "Exact Learning Augmented Naïve Bayes Classifier," *Entropy*, vol. 23, is. 12, no. 1703, Dec. 2021, doi: 10.3390/e23121703.
- [8] L. Nguyen, "Tutorial on Support Vector Machine," *Applied and Computational Mathematics*, vol. 6, is. 4, pp. 1–15, Jul. 2017, doi: 10.11648/j.acm.s.2017060401.11.
- [9] M. Popovac, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, "Convolutional Neural Network Based SMS Spam Detection," *26th Telecommunications Forum, TELFOR 2018 - Proceedings*, 18395017, Nov. 20-21 2018, doi: 10.1109/TELFOR.2018.8611916.
- [10] N. Liubchenko, A. Podorozhniak, and V. Bondarchuk, "Neural network method of intellectual processing of multispectral images," *Advanced Information Systems*, vol. 1, is. 2, pp. 39–44, Jan. 2017, doi: 10.20998/2522-9052.2017.2.07.
- [11] *SMS Spam Collection Dataset*, Kaggle Datasets, Oct. 2022. [Online]. Available: <https://www.kaggle.com/uciml/sms-spam-collection-dataset>.
- [12] *Spam Mails Dataset*, Kaggle Datasets, Oct. 2022. [Online]. Available: <https://www.kaggle.com/datasets/venky73/spam-mails-dataset>.

Проектування виконавчого пристрою системи контролю управління доступу на базі STM32F103

Олександр Жеребцов, Арсеній Рассохін

*Кафедра інформаційних технологій
Запорізький інститут економіки та інформаційних технологій
Запоріжжя, Україна*

Анотація—Розглянуто актуальні методи побудови систем контролю та управління доступом у побутових та промислових випадках застосування. Обґрунтовано рішення розробити і запровадити клієнт-серверну систему контролю та управління доступом (СКУД), основою якої є, керуючий периферійними пристроями, однокристальний комп'ютер, який підключено до локальної мережі. Описано програмно-апаратні засоби реалізації виконавчого вузлу СКУД, здійснено розробку принципової електричної схеми та виготовлено друковану плату пристрою.

Ключові слова—Комп'ютерна інженерія, СКУД, друкована плата, STM32F103.

I. ВСТУП

За визначенням «Вікіпедії» СКУД це комплекс технічних та програмних засобів безпеки, що здійснює регулювання входу / виходу та переміщень людей чи транспортних об'єктів на територіях, які знаходяться під охороною, для адміністративного моніторингу та попереджень несанкціонованого проникнення [1].

На даний час вже існує безліч реалізацій систем контролю, що давно вийшли за межі їх базової функції- надання доступом до чогось або до когось за наявністю прав [2].

Світовий ринок подібних систем здатен задовільнити будь-якого прискіпливого замовника, та навіть більше, демонструючи різноманітні цінові та якісні пропозиції, починаючи від бюджетних пристроїв локального використання, де не потрібен високий фізичний та програмний рівень захисту до доволі недешевих клієнт-серверних систем, що спроектовані безперервно охороняти фінансові заклади або об'єкти стратегічного значення [3].

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Системи контролю управління доступом поділяються на два базових різновиди за системною топологією, від яких надалі відбувається розширення властивостей: автономні та мережеві. Також дані системи можна класифікувати за кількістю пунктів пропуску, різновидом пристрою зчитування (біометричні, цифрові, кодові), за різновидом виконавчого пристрою, рівнем фізичного захисту, методом мережевої комунікації [4].

Виконавчий пристрій системи контролю - це фізичний механізм, що призначений обмежувати доступ людини або предмету на територію, що пристрій охороняє [5].

RFID (Radio frequency identification) або радіо-частотна ідентифікація супроводжує історію ще з 1937 року, коли воєнні авіаційні системи навігації вперше використали спосіб розпізнавання «свій-чужий». Вона активно використовувалася під час другої світової війни. Перша демонстрація сучасних RFID-мікросхем як активного так і пасивного типу датується 1973 роком. Портативна система працювала на частоті 915МГц та використовувала 12-

розрядні мітки. Аж у 1999 році була створена «глобальна система стандартів» для мікросхем RFID [6].

Однією з найбільш популярних компаній, що спеціалізується на проектуванні та виробництві систем контролю та управління доступом є «Bosch». Німецький гігант розробляє велику кількість різновидів СКУД, різного рівня захисту, екстраполяції та розгалуження, починаючи від локальних мережевих систем контролю, закінчуючи потужними контролерами для великих підприємств та установ. У якості аналогу розробки можна виділити «Access Easy Controller 2.1 (AEC2.1)» від «Bosch Security System»- це нове покоління контролера доступу, який поєднує в собі можливості вбудованого веб-сервера та системи відеоспостереження [7].

Здійснивши класифікацію й огляд СКУД можна створити загальну схему проекту (рис. 1), відокремити блоки, підібрати протоколи зв'язку, за якими вони будуть взаємодіяти.

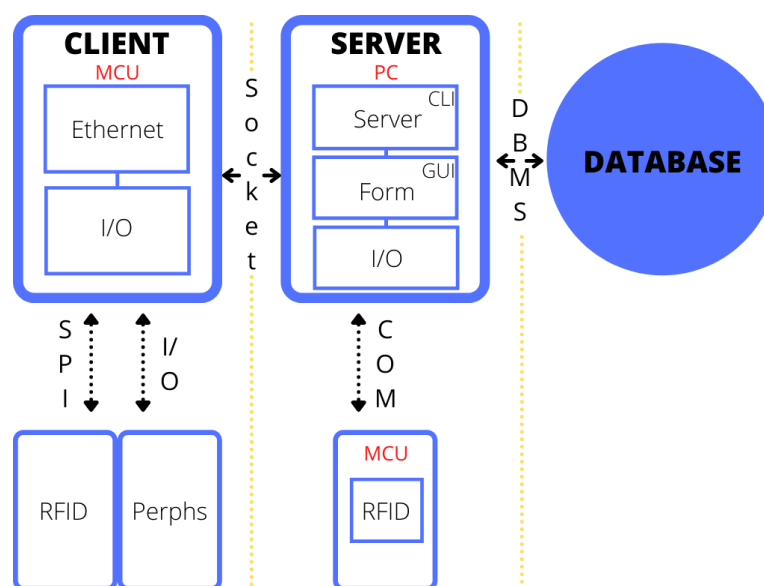


Рисунок 1 – Загальна схема системи

Структура системи запозичує основу клієнт-серверної топології будовання подібних комплексів. Як було вказано, принцип побудови такого зразка надає більш широкий функціонал під час роботи та має вищий рівень безпеки опрацювання даних [8].

Система має керуватись із спеціального пункту контролю, який має знаходитись на деякому віддаленні від основного «шлюзу», де реалізується фізичний пропуск. За допомогою спеціального пункту контролю мають бути реалізовані такі можливості як: зберігання даних користувачів та їх унікальних ідентифікаційних номерів, безперервний контроль та фіксація дати та часу подій входу та виходу користувачів, внесення, зміна та видалення даних користувачів.

СКУД має відрізнятись властивостями до боротьби із електромагнітними, електричними та несанаційними програмними намаганнями втрутитись у її робочий процес. Також, має бути витриманий баланс між бюджетними прототипами локальних систем контролю та професійними реалізаціями клієнт-серверних систем, які працюють на підприємствах та у деяких навчальних закладах, тощо.

Отже, розроблена система контролю та управління доступом належить до систем мережевого клієнт-серверного типу, з одним пунктом пропуску, але з апаратною та програмною можливістю розширювати комплекс, з'єднуючи в масив. Спосіб ідентифікації-радіочастотний з унікальним номером ідентифікації «UID».

III. МЕТОДИ

Основою реалізації було обрано мікроконтролер STM32F103. Потужний мікроконтролер із обчислювальним ядром мікропроцесора Cortex-M3 з частотою обчислювання ядра та головної шини 72МГц. Також присутня велика кількість портів вводу-виводу (I/O ports) та два окремих незалежних SPI (Serial Peripheral Interface) протоколи.

У якості вузлу, що забезпечить мережеве з'єднання, було обрано WIZnet W5500- це вбудований Ethernet-контролер з підключенням TCP/IP, який забезпечує спрощене підключення до інтернету або локальної мережі для вбудованих систем за допомогою SPI (Serial Peripheral Interface).

W5500 найкраще підходить у випадках, коли потребується стабільне підключення до інтернету або локальної мережі, використовуючи єдиний чіп для реалізації стека TCP/IP, 10/100 Ethernet MAC та PHY. Стек TCP/IP підтримує TCP, UDP, IPv4, ICMP, ARP, IGMP і PPPoE та інші. Це було доведено в різних програмах протягом багатьох років. W5500 використовує внутрішній буфер 32 Кбайт як свою пам'ять для передачі даних.

Для проектування принципової електричної схеми, друкованої плати було обрано EasyEda - крос-платформне веб-орієнтоване середовище автоматизації проектування електроніки, що включає редактор принципів схем, редактор топології друкованих плат, SPICE-симулятор, хмарне сховище даних, систему управління проектами, а також засоби підготовки та замовлення виготовлення друкованих плат.

IV. РЕЗУЛЬТАТИ

Використовуючи програмно-апаратний стек розробки, було створено виконавчий вузол СКУД (рис. 2), його принципову електричну схему (рис. 3) на базі мікроконтролера STM32F103.

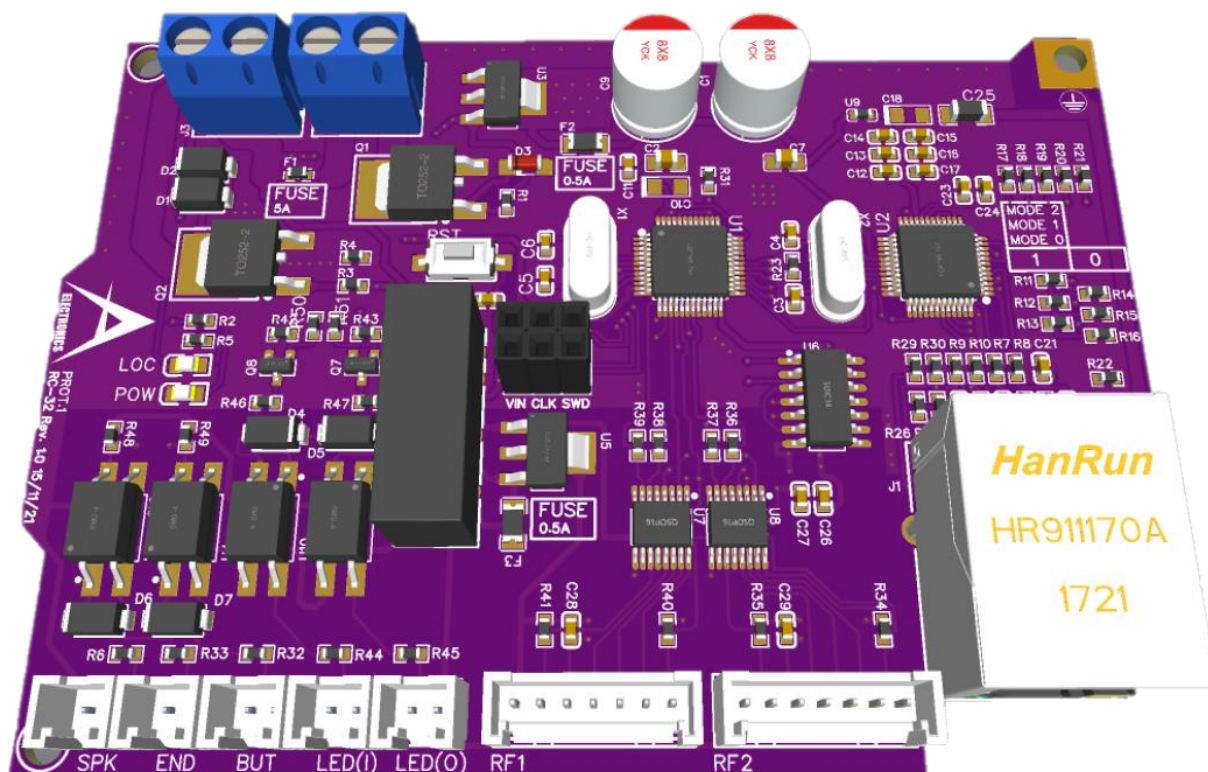


Рисунок 2 - Зображення моделі друкованої плати виконавчого вузлу

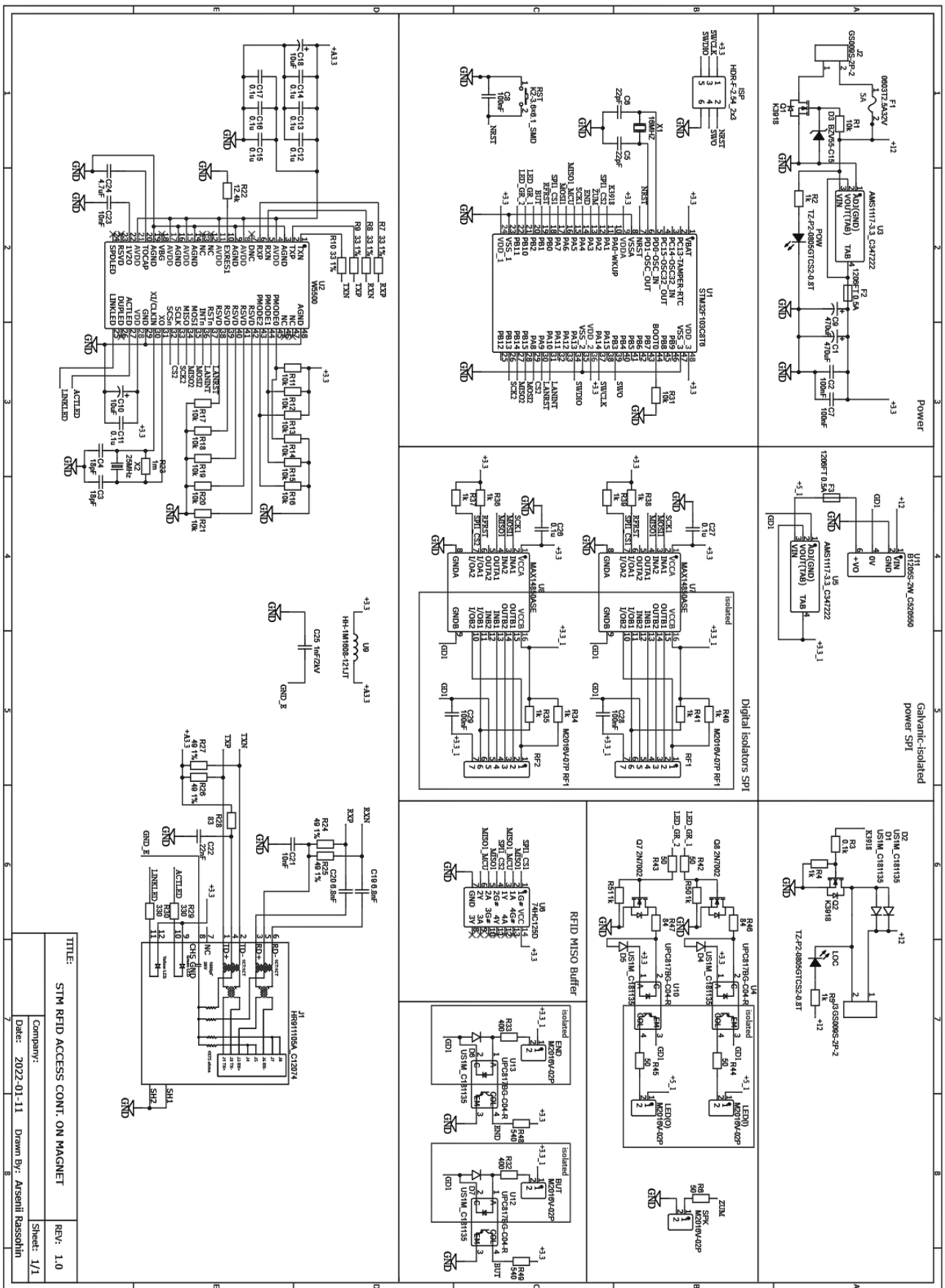


Рисунок 3 – Принципова електрична схема виконавчого вузлу
 (Schematic_STM32F103-RFID-D3MG_2022-11-17.png)

V. ОБГОВОРЕННЯ

Перш за все, треба проаналізувати умови використання системи, яка є системою захисту та контролю. Перед тим, як вона почне процес контролю, її треба захистити саму. Насамперед, як і будь-який електронний прилад, що має велику кількість електронних напівпровідників та комунікує із іншим приладом методом фізичного поєднання їх струмопровідних провідників, прилад «боїться» впливу електричного струму та електромагнітних перешкод.

Через це, обов'язковими є вузли захисту від зовнішніх впливів, оптичні, що мають діелектричний проміжок. Почнемо з вузлів з'єднання електронних приладів безконтактного зчитування MFRC 522 із головним мікроконтролером STM32F103, для чого були обрані мікросхеми MAX 14850, які здатні «розв'язати», за допомогою оптичного бар'єру, інформаційні шини мікроконтролера із зовнішнім впливом (рис. 4).

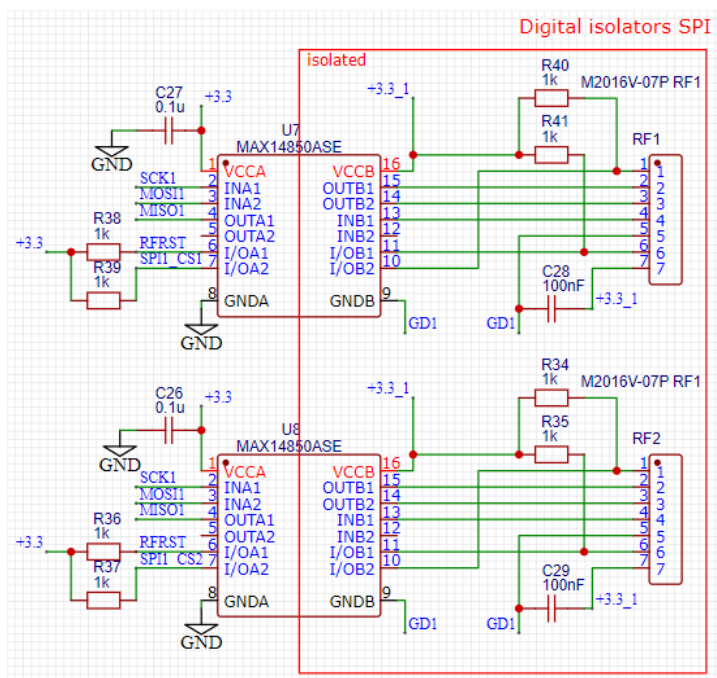


Рисунок 4 – Зображення частини діелектричного захисту

Також виконавчий пристрій має органи керування: кнопки, індикацію та кінцевий вимикач. Вони також безпосередньо контактують із периферійними контактами мікроконтролера, де будь-який несанкціонований вплив на них може вивести мікросхему із робочого стану. Щоб уникнути цього- використовуються оптичні пари PC 817 (рис. 5). Вони також мають бар'єр від струмопровідних частин на зовні.

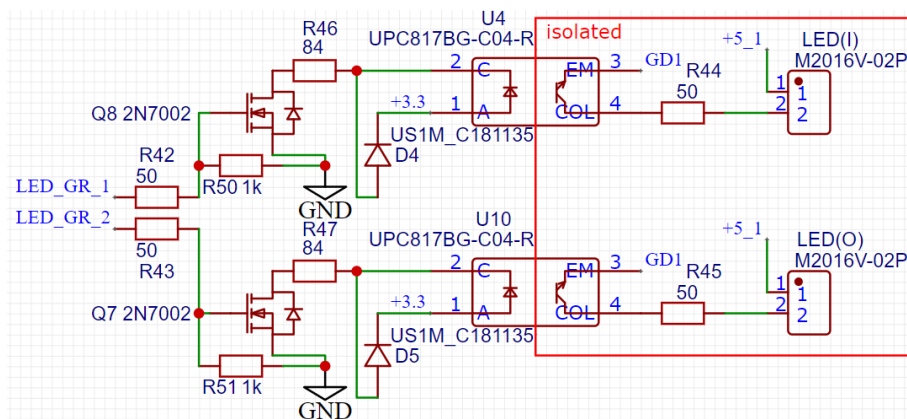


Рисунок 5 – Зображення частини діелектричного захисту

Буферна мікросхема 74HC125D (рис. 6) запобігає несвоєчасному спрацюванню електронних приладів безконтактного зчитування на мікросхемі MFRC 522. Виключає можливість виникнення помилок на інформаційній шині SPI. За допомогою неї відбувається «спілкування» зчитувачів із основним мікроконтролером.

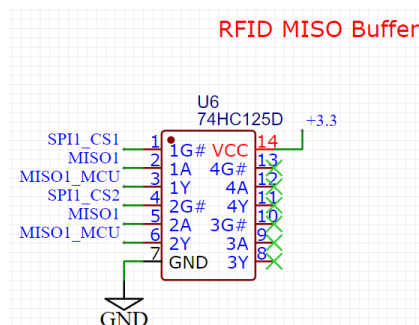


Рисунок 6 – Зображення частини буферної мікросхеми

Живиться схема від напруги 12 вольт, через те, що більшість пристроїв фізичного запобігання проникненню працюють від цієї напруги, вона у більшості випадків є універсальною. Сама ж логічна частина схеми, починаючи від мікроконтролера, закінчуючи контролером локальної мережі живиться від напруги логічного рівня 3,3 вольта. Для цього потрібен вузол стабілізації напруги. Його параметри мають відповідати нормам низької пульсації постійної напруги для запобігання потрапляння перешкод на лінію живлення, що допоможе уникнути можливих проблем із працездатністю обчислювальних пристроїв та передавачів.

Вузол перетворення та стабілізації напруги побудований на мікросхемі AMS 1117 представлено на рисунку 7. Перед вузлом стабілізації, згідно стандартам, присутня схема запобігання так званої «переполносовки» живлення, коли при підключення плутаються провідники за полюсами. Також, розташовано силовий запобіжник для захисту сильно-струмової частини схеми та логічний запобіжник, що має захистити схему обчислювальних блоків у випадку виникнення нестандартних ситуацій.

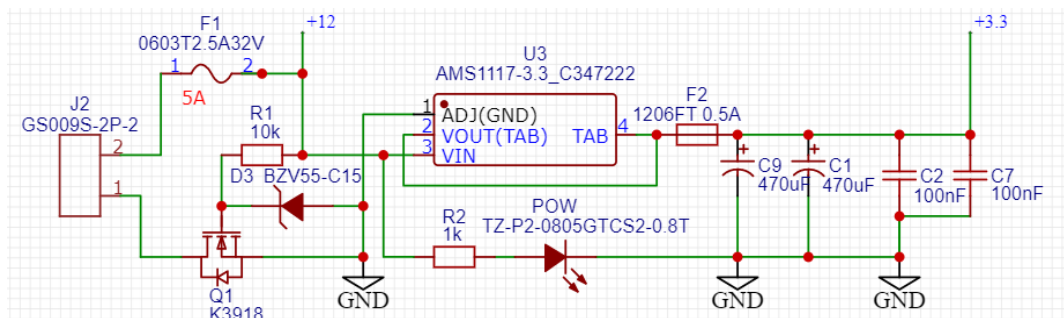
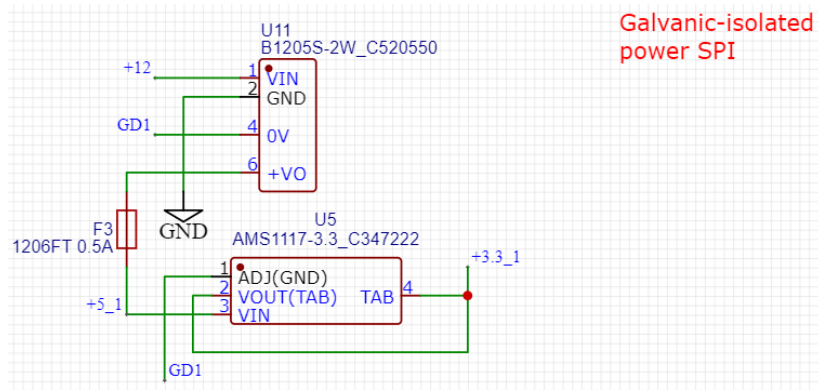


Рисунок 7 – Зображення частини блоку живлення контролера

Звісно ж, захист живлення присутній і на приладах безконтактного зчитування, що розташовані не на друкованій платі виконавчого пристрою, а зовні. В якості ізолятора живлення- ізоляційний перетворювач імпульсно-трансформаторного типу з гальванічною «розв'язкою» виходу В1205, стабілізатор AMS 1117 (рис. 8).

Вузол комутації зовнішнього навантаження, а саме електричного магніту, під який проектувався пристрій заснований на силовому транзисторі MOSFET-типу K3918, що здатен до комутації великого струму (рис. 9).



Galvanic-isolated
power SPI

Рисунок 8 – Зображення частини блоку живлення периферії

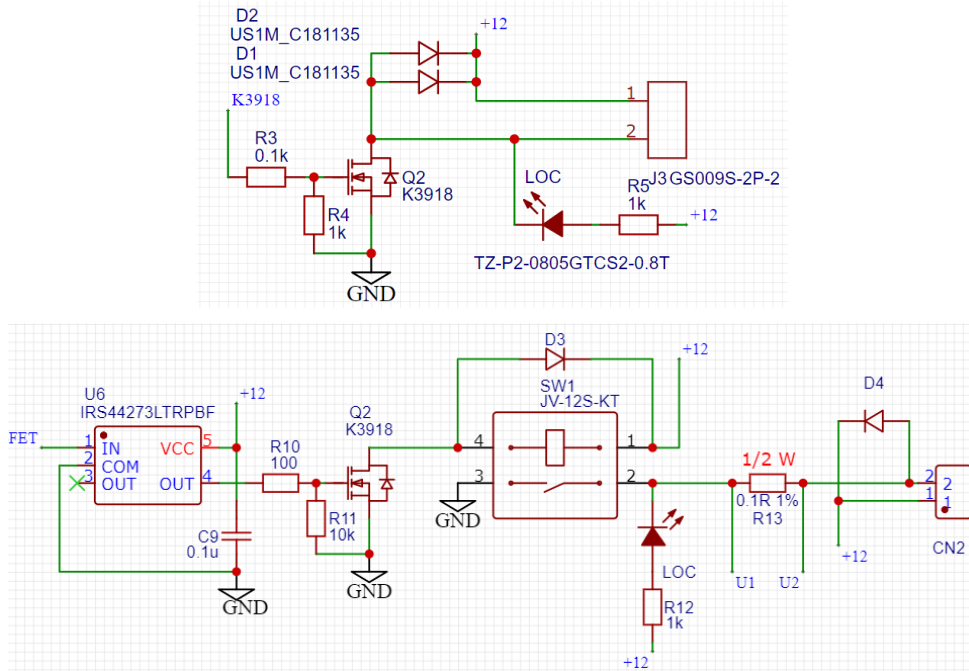


Рисунок 9 – Зображення частини комутаційного вузла

Частину з системою вбудованого Ethernet-контролеру WIZnet W5500 наведено на рисунку 10. Особлива увага приділена до якості вузла живлення контролера. Мається роз'єднувальний трансформатор, що ізолює з'єднання Ethernet фізичного протоколу.

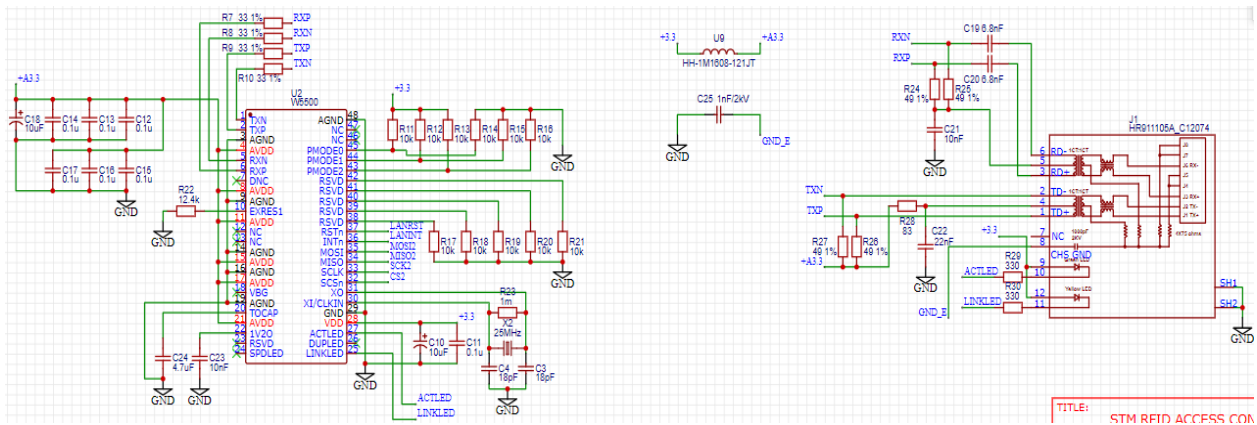


Рисунок 10 – Зображення частини блоку Ethernet-контролеру

STM32F103 потужний мікроконтролер із обчислювальним ядром мікропроцесора Cortex-M3 з частотою обчислювання ядра та головної шини 72МГц, який є центральним вузлом обчислення для всіх периферійних пристроїв. Тактування відбувається зовнішнім резонатором кварцового типу (рис. 11).

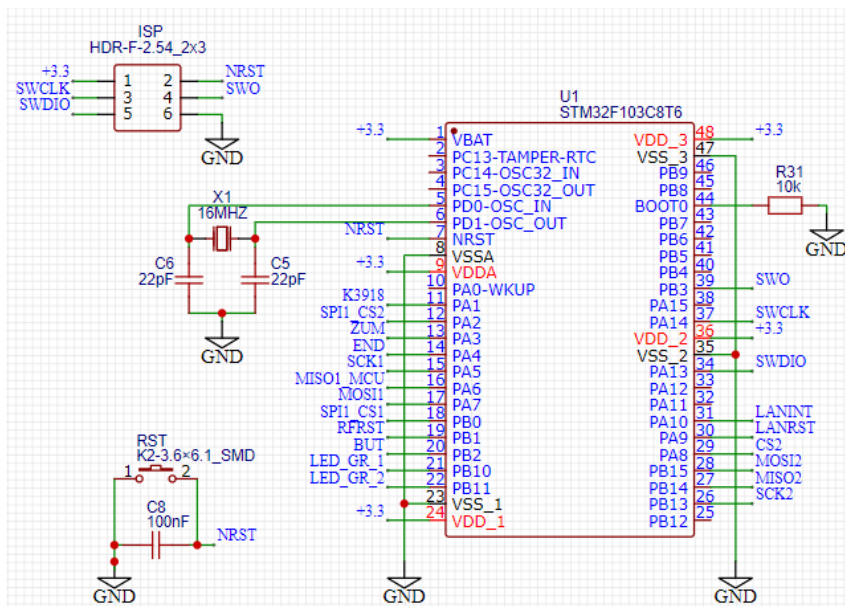


Рисунок 11 – Зображення частини головного обчислювального блоку

VI. ВИСНОВКИ

В рамках створеної СКУД було розроблено прилад виконавчого пристрою на базі однокристалного ЕОП, з додаванням фізичного контролеру безпечного мережевого зв'язку. Даний пристрій володіє достатнім функціоналом для забезпечення надійного користування у СКУД та рекомендовано до практичного застосування.

ЛІТЕРАТУРА

- [1] Система контролю і управління доступом, Wikipedia, 2022. Accessed on: Sept. 19, 2022 [Online]. Available: https://uk.wikipedia.org/wiki/Система_контролю_та_управління_доступом.
- [2] M. Benantar, *Access Control Systems: Security, Identity Management and Trust Models*, New York, NY, USA: Springer, 2006, doi: 10.1007/0-387-27716-1
- [3] M. Broy, K. Stølen, "Access Control System. In: Specification and Development of Interactive Systems," *Monographs in Computer Science*, New York, NY, USA: Springer, 2001, doi: 10.1007/978-1-4613-0091-5_9
- [4] T. Hof, E. Robert, I. Barthes, S. Bastard., "Design and implementation of a Decentralized Access Control System (DACs) and an application," in *New Technologies, Mobility and Security*, H. Labiod, M. Badra, Eds., Paris, France: Springer Dordrecht, 2007, pp. 621-623, doi:10.1007/978-1-4020-6270-4_53
- [5] IP Door Access Control., Kintroniks IP Security solutions, 2022. Accessed on: Sept. 19, 2022 [Online]. Available: <https://kintroniks.com/solutions/ip-door-access-control/>
- [6] J. Oxeer, H. Blemings, "RFID Access Control System," in: *Practical Arduino*. Berkeley, CA, USA: Apress Berkeley, 2009, pp. 273-299, doi:10.1007/978-1-4302-2478-5_14
- [7] Bosch, Gerlingen, Germany, *Access Easy Controller 2.1. v.3* (2010), Accessed on: Sept. 19, 2022 [Online]. Available: https://resources-boschsecurity-cdn.azureedge.net/public/documents/AEC_Data_sheet_enUS
- [8] А. Рассохін, О. Жеребцов, "Розробка автоматизованої системи контролю управління доступом," *матеріали Тринадцятої Всеукраїнської, Двадцятої регіональної наукової конференції молодих дослідників «Актуальні проблеми математики та інформатики»*, ЗНУ, Запоріжжя, Україна: Видавництво ЗНУ, 2022, с. 43

Застосування нечітких алгоритмів в задачах розробки інформаційних систем для прогнозування результатів кіберспортивних змагань

Юлія Казимиренко, Оксана Гайдаєнко,
Владислав Принць, Анастасія Байталюк

*Кафедра інформаційних управляючих систем і технологій
Національний університет кораблебудування імені адмірала Макарова
Миколаїв, Україна*

Анотація—Розглянуто принципи інформаційної підтримки кіберспортивних змагань. В основу створення інформаційної системи покладено новий алгоритм, розроблений за допомогою методів нечітких множин і нечітких відношень. Результати роботи спрямовані на вирішення важливого науково-прикладного завдання інформаційного забезпечення систематизації та прогнозування результатів кіберспортивних ігор.

Ключові слова—*Інформаційні технології, Кіберспорт, Ймовірнісні методи, Експертні оцінки, Раціональні компромісні рішення.*

I. ВСТУП

У сучасному суспільстві поступово відбувається переоцінювання важливості кіберспортивних ігор, розвиток яких пройшов тривалий шлях від «азартних» ігор до постановки і розв'язання складних математичних задач з прогнозуванням очікуваних результатів у режимі реального часу. Головним проблемним питанням слід вважати непередбачуваність подій з невизначеними зв'язками у системі.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Теоретичним підґрунтям з математичного моделювання букмекерської індустрії є теорія ігор, яку автори роботи [1] розглядають як «науку про стратегічне мислення». В роботі [2, с. 72-79] математичні моделі теорії ігор мають вигляд досліджень процесів прийняття рішень в умовах пошуку компромісних рішень. В роботі [3, с. 43-49] розглянуто ускладнену постановку задачі: додаються умови невизначеності з аналітичним розвитком подій. Проте слід розуміти, що очікувані результати з кожної події є наслідком збігу різних факторів зі змінним невизначеним впливом. Їх логічний та статистичний аналіз доцільно покласти в основу проектування інформаційної системи, використання якої допоможе аналізувати ситуації з метою виключення негативних факторів.

III. МЕТОДИ

Методологічною основою дослідження є теорія нечітких множин та нечітких відношень [5, с. 113-150], що дає змогу досліджувати поставлені задачі у динаміці з оцінкою вихідних значень. В основу дослідження покладено гіпотезу про мінімізацію помилки гравця. Відповідно до ситуації прийняття рішень у формі звичайного відношення переваги, пошук раціонального рішення полягатиме у виборі максимальних (недомінованих) альтернатив, визначення яких відбувається за результатами експертних оцінок минулих матчів. В основу

функціонування інформаційної системи покладено вимоги надання обробленої статистики користувачеві; аналіз внутрішніх ігрових факторів; отримання історії та статистики особистих зустрічей команди та виконання прогнозів внутрішніх ігрових показників. Реалізація апарату здійснюється з використанням нечітких нейронних мереж (fuzzy-neural networks), для підбору параметрів застосовується метод зворотного розповсюдження помилки.

IV. РЕЗУЛЬТАТИ

Головною проблемою у прогнозуванні результатів кіберспортивних ігор є непередбачуваність та неочевидність зав'язків у системі, тому в роботі введено новий єдиний фактор – фіксоване число перемог, що робить процес оптимізації однокритеріальним. Для проектування інформаційної системи розроблено алгоритм оптимального пошуку з «фіксованим шансом на перемогу». Для цього на підставі статистичної обробки власного досвіду гравців складається шкала, на яку накладено події з виграшу. Потім розраховується приблизна кількісна оцінка з наступним уточненням. Завдяки накладанню на шкалу фактичних даних можна розраховувати «фіксований шанс на перемогу» для певної кількості гравців. Розроблений алгоритм являє собою теоретичну складову проекту інформаційної системи та не враховує змагальну та емоційну складові кіберспортивних змагань, які, у свою чергу, сприятимуть формуванню некоректної бази для фінансового і маркетингового аналізу. Інформаційна система взаємодіє з користувачем через web-сайт букмекерської контори.

V. ОБГОВОРЕННЯ

Результати роботи спрямовані на вирішення важливого науково-прикладного завдання інформаційного забезпечення систематизації та прогнозування результатів кіберспортивних ігор. Застосування методу нечітких множин і нечітких відношень сприятиме підвищенню ефективності вирішення фінансових і маркетингових задач з пошуком раціональних компромісних рішень.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Перспективи подальших досліджень пов'язані з тимчасовою (тобто для певного сезону змагань) апробацією розробленого алгоритму у студентському середовищі кіберспортивних змагань, що дає змогу виявити неточності у його роботі та усунути недоліки.

VII. ВИСНОВКИ

За допомогою методу нечітких множин і нечітких відношень розроблено алгоритм раціонального пошуку з «фіксованим шансом на перемогу», який покладено в основу створення нової інформаційної системи для прогнозування результатів кіберспортивних змагань.

ЛІТЕРАТУРА

- [1] R. J. Aumann, *Lectures on Game Theory*. San Francisco, USA: Westview Press, 1989.
- [2] С. А. Ус, та Л. С. Корішкіна. *Моделі і методи прийняття рішень*. Дніпропетровськ, Україна: НГУ, 2014.
- [3] А. А. Шиян, *Теорія ігор: основи та застосування в економіці та менеджменті*. Вінниця, Україна: ВНТУ, 2009.
- [4] І. О. Лазнева, Д. І. Цараненко, «Кіберспорт та його вплив на зміну структури світового ринку комп'ютерних ігор», *Науковий вісник Ужгородського національного університету*, № 22(2), с. 63-67, 2018.
- [5] Є. А. Лавров, Л. П. Перхун, В. В. Щендрик, Е. Г. Кузнецов, Парафененко Ю. В., Сергієнко В. А., *Математичні методи дослідження операцій*. Суми, Україна: Сум.ДУ, 2017.

(ПЕРЕЛІК ЛІТЕРАТУРИ СЛІД ПОДАВАТИ У СТИЛІ IEEE)

Реалізація алгоритму швидкого перетворення Фур'є на ПЛІС

Андрій Павлишин, Володимир Грига, Володимир Мандзюк

*Кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація — У роботі розглянуто алгоритм дискретного перетворення Фур'є та його більш ефективний підвид – швидке перетворення Фур'є. Описано апаратну реалізацію швидкого перетворення Фур'є, структуру пристрою та його компонентів, процес обробки та принципи виконання обчислень. Проведено функціональне моделювання алгоритму швидкого перетворення Фур'є та отримано апаратні затрати та часові параметри при його синтезі на ПЛІС сімейства Artix-7 фірми Xilinx з використанням системи автоматизованого проектування Vivado IDE.

Ключові слова— *Обробка сигналів, Алгоритм, Програмовані логічні інтегральні схеми, Швидке перетворення Фур'є, FPGA-пристрій, Апаратні засоби, Програмні засоби.*

І.ВСТУП

Серед усіх ортогональних тригонометричних перетворень (ОТП) універсальністю відзначається перетворення Фур'є. На його основі можуть бути легко виражені такі фундаментальні операції над сигналами як згортка, диференціювання та інтегрування, описано багато проблемних процедур обробки сигналів, а інші перетворення, як наприклад Френеля, Гільберта, Лапласа, обчислені на його основі. Часткові випадки перетворення Фур'є, такі як синус- та косинус перетворення Фур'є, використовуються в задачах кодування даних [1].

Серед існуючих методів отримання швидких алгоритмів завдяки своїй простоті найбільше розповсюдження отримав метод Кулі-Тьюкі. У випадку дискретного перетворення Фур'є в отриманих за методом Кулі-Тьюкі алгоритмах швидкого перетворення Фур'є (ШПФ) використовуються комплексні фазові множники, що вимагає значних затрат для реалізації постійного запам'ятовуючого пристрою збереження фазових множників, а також для реалізації базової операції. Уперше алгоритм ШПФ, в якому фазові множники мають структуру дійсних чисел, був отриманий Рейдером і Бреннером.

Реалізація ШПФ на апаратній основі (цифрова логіка, FPGA-пристрої) має перевагу з точки зору високошвидкісного перетворення в реальному часі, але є дещо складнішою ніж програмна реалізація алгоритму. Програмна реалізація алгоритму виконує інструкції послідовно (одна за одною) і тому сильно обмежена пропускну здатністю процесора. Апаратне ШПФ здійснює багато задач паралельно, тому може досягти на порядок більшу пропускну здатність, ніж програма, яка виконується на мікропроцесорі. Проте алгоритм ШПФ при реалізації в апаратному вигляді потребує використання додаткових засобів, щоб утримувати розмір та швидкість логіки в межах практичного використання. Тому метою даної публікації є апаратна реалізація структури логічного модуля, який виконує швидке перетворення Фур'є та аналіз його системних характеристик при синтезі на FPGA-пристроях фірми Xilinx.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Зменшення технологічних норм для виготовлення пристроїв напівпровідникової інтегральної мікро- та наноелектроніки призводить до ускладнення їх виготовлення і в багатьох випадках не дає значного приросту продуктивності. Все частіше підвищення продуктивності комп'ютерних систем здійснюється за допомогою додавання спеціалізованих блоків, які мають більшу ефективність за рахунок паралельного виконання та конвеєризації операцій [1]. Програмовані логічні інтегральні схеми (ПЛІС), зокрема FPGA-пристрої, які дозволяють реалізовувати та виконувати функції апаратних структур, стають все доступнішими та мають широку затребуваність серед розробників. Це, у свою чергу, дає можливість проектувати системи з апаратним виконанням функцій, замість реалізації на мікропроцесорах загального призначення [2].

На сьогоднішній день для реалізації ШПФ застосовуються програмні засоби, коли процесор чи мікроконтролер загального призначення згідно написаної програми виконує обробку сигналу. При виконанні одного типу операцій над великим потоком даних цей підхід не є найефективнішим, оскільки швидкодія тут обмежена тим, наскільки швидко процесор покроково виконує всі операції алгоритму. У цьому випадку спеціалізовані апаратні модулі або програмовані логічні пристрої, завдяки можливості паралельного виконання операцій та конвеєризації, мають суттєво більшу швидкодію.

Один із варіантів структури апаратного процесора, яка використовується для реалізації ШПФ, зображений на рис. 1 [3].

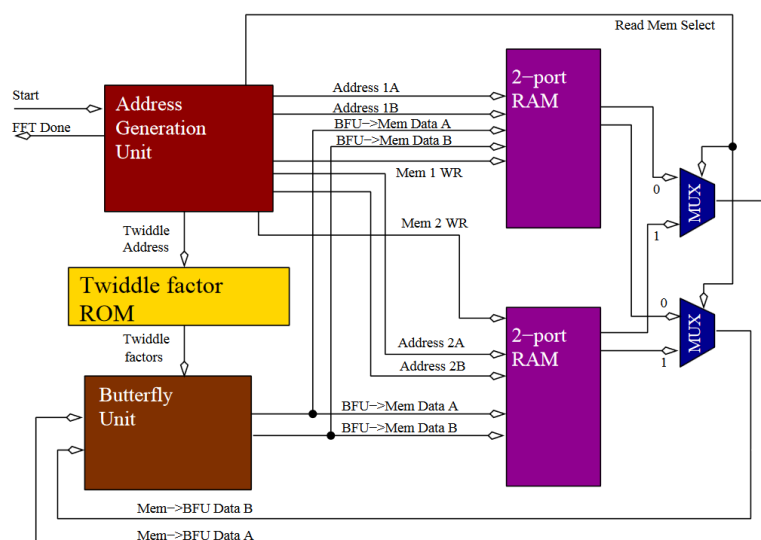


Рис. 1. Загальна структура апаратного процесора ШПФ [3].

На даному рисунку блок Address generation unit (AGU) являє собою модуль генерування адрес, який контролює генерування адрес для зчитування та запису вмісту пам'яті модуля виконання операції "метелик" (Butterfly unit). AGU також генерує сигнали, які контролюють записи в пам'ять, а також, з якого модуля пам'яті здійснюється зчитування [3]. У цій структурі використовуються два модулі двопортової оперативної (2-port RAM) пам'яті. Усі шини даних, показані на діаграмі, здійснюють перенесення дійсної та уявної (комплексної) частини даних (для передачі реальної і уявної частини даних необхідна подвійна розрядність всіх шин).

Для реалізації алгоритму ШПФ з можливістю застосування FPGA-пристрою в роботі використано алгоритм Кулі-Тьюкі [4], який є дуже ефективним, найлегшим для реалізації та широко використовуваний на практиці.

Швидке перетворення Фур'є є різновидом дискретного перетворення Фур'є (ДПФ).
Дискретне перетворення визначається як:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-j\frac{2\pi k}{N}n},$$

де послідовність $\{x_n\} = x_0, x_1, \dots, x_{N-1}$ – вибірка даних, а $\{X_k\} = X_0, X_1, \dots, X_{N-1}$ – результат перетворення.

Якщо взяти, що комплексні коефіцієнти $e^{-j\frac{2\pi k}{N}n}$ є попередньо розраховані та зберігаються в пам'яті, то для обчислення одного X_k потрібно виконати N комплексних множень та $(N-1)$ комплексних додавань. Тобто, для цілого перетворення Фур'є потрібно $(2N^2-N)$ обчислень, при достатньо великих N основною складовою тут буде N^2 . Через цю велику складність алгоритму в більшості випадків використовують ШПФ.

Алгоритм ШПФ використовує властивості експонентної константи $e^{-j\frac{2\pi k}{N}kn}$. Для зручності введемо символ, який буде позначати експонентну константу $W_N = e^{-j\frac{2\pi}{N}}$.

Запишемо дві властивості, які тут використовуються:

- 1) симетрія комплексно-спряжених чисел: $W_N^{k(N-n)} = W_N^{-kn} = (W_N^{kn})^*$;
- 2) періодичність експоненти: $W_N^{kn} = W_N^{k(N+n)} = W_N^{(k+N)n}$.

Обидві властивості виходять з того, що $e^{-j2\pi k} = 1$, тобто якщо W_N має в степені N , то знаменник скорочується і константа стає одиницею.

Крім цього, в ШПФ використовується розбиття одного перетворення Фур'є на два менших. Зокрема, сумування ДПФ може бути поділене наступним чином:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot W_N^{kn} = \sum_{n=0}^{N/2-1} x_{2n} \cdot W_N^{k2n} + \sum_{n=0}^{N/2-1} x_{2n+1} \cdot W_N^{k(2n+1)}.$$

Тут сума всіх членів перетворення поділена на суму парних членів перетворення та суму непарних членів перетворення. Далі можна винести з степенів W частини, які не залежать від n .

$$\begin{aligned} X_k &= \sum_{n=0}^{N/2-1} x_{2n} \cdot W_N^{k2n} + \sum_{n=0}^{N/2-1} x_{2n+1} \cdot W_N^{k(2n+1)} = \\ &= \sum_{n=0}^{N/2-1} x_{2n} \cdot (W_N^2)^{kn} + W_N^k \cdot \sum_{n=0}^{N/2-1} x_{2n+1} \cdot (W_N^2)^{kn}. \end{aligned}$$

У цій формулі можна переписати W_N^2 як $W_{N/2}$, тому що

$$W_N^2 = e^{-j\frac{2\pi}{N}2} = e^{-j\frac{2\pi}{N/2}} = W_{N/2}.$$

Після цього формула виглядатиме наступним чином:

$$X_k = \sum_{n=0}^{N/2-1} x_{2n} \cdot W_{N/2}^{kn} + \sum_{n=0}^{N/2-1} x_{2n+1} \cdot W_{N/2}^{kn}.$$

У результаті отримаємо, що перший доданок – це ДПФ парних вибірок початкової послідовності x , а другий доданок – ДПФ непарних вибірок з x .

Отже, ДПФ можна розкласти як суму з двох в два рази менших ДПФ.

Нижче наведено графічне представлення розбиття ДПФ розміром $N = 8$ (рис. 2).

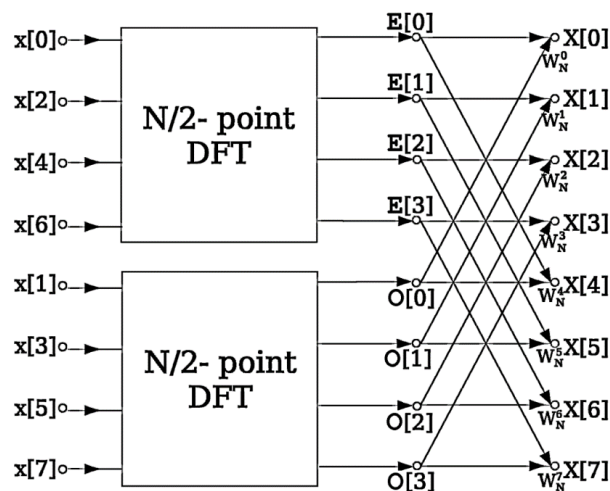


Рис. 2. Структура перетворення Фур'є, яке складається з двох менших перетворень.

У верхній частині рисунку показано парні вибірки: $x[0]$, $x[2]$, $x[4]$, $x[6]$, які йдуть на вхід 4-х точкового ДПФ, яке на виході продукує коефіцієнти, позначені E . У нижній частині рисунку зібрані непарні вибірки: $x[1]$, $x[3]$, $x[5]$, $x[7]$, які йдуть на вхід іншого 4-х точкового ДПФ, яке на виході продукує коефіцієнти, позначені O .

На наступному етапі здійснюється обчислення коефіцієнтів повного перетворення з використанням вихідних коефіцієнтів E та O з двох менших ДПФ. На діаграмі видно, щоб отримати $X[0]$ потрібно до $E[0]$ додати добуток W_8^0 і $O[0]$. Для отримання $X[7]$ формула буде наступною: $X[7] = E[3] + W_8^7 \cdot O[3]$. Це і є загальна формула обчислення коефіцієнтів загального перетворення з використанням коефіцієнтів парного та непарного перетворень.

Якщо розглянути обчислювальну складність перетворення з таким поділом, отримаємо два стандартні ДПФ розміром $N/2$ та N додавань та N мнужень на коефіцієнти W з різними степенями. Коефіцієнти W можна один раз попередньо обчислити і зберігати в пам'яті для подальшого використання.

При розгляді операції множення отримаємо:

$$\left(\frac{N}{2}\right)^2 \cdot 2 + N \approx \frac{N^2}{2} + N.$$

При дуже великому N ця оцінка може бути прирівняна до $N^2/2$. Тобто, в результаті поділу перетворення на два менших перетворення та комбінування їх виходів зменшило кількість необхідних операцій в двічі.

Поділ можна проводити, доки не дійдемо до одноточкових перетворень Фур'є. В кінці отримується необхідна кількість операцій, рівна $N \log_2 N$, що і являє собою ШПФ.

На рис. 3 показано графічне представлення ШПФ розміром $N = 8 = 2^3$, в якому присутні три етапи обчислень: два проміжні і один кінцевий.

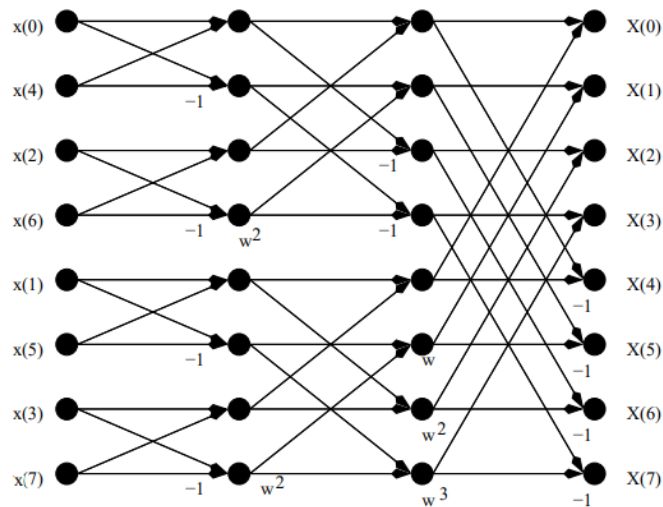


Рис. 3. Графічне представлення обчислення ШПФ з розміром 8 точок.

Як слідує з рис. 3, базовим обчисленням, яке виконується у всьому процесі, є двоштокове перетворення Фур'є, яке називають "метеликом".

Коли виконується поділ перетворення на парні та непарні вибірки, а потім ці блоки теж діляться на парні та непарні елементи, то на вході отримується певна послідовність вибірок. Для $N = 8$ це будуть вибірки з індексами 0, 4, 2, 6, 1, 5, 3, 7. Виявляється, якщо взяти позиції цих вибірок в послідовності у двійковому вигляді та обернути порядок бітів, то отримаємо цю послідовність. Наприклад, елемент 0 має двійкове представлення 000, оскільки тут лише нулі, то у зворотному порядку матимемо теж 0. Елемент 1 у двійковому вигляді буде 001, а після обернення порядку бітів $100 = 4$. Двійка у двійковому вигляді симетрична 010, тому вона теж залишається як є. Трійка буде 011, а після обернення $110 = 6$. І так далі.

III. РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

При апаратній реалізації ШПФ є багато значних відмінностей на відміну від програмної реалізації. Найголовнішим є те, що апаратне перетворення Фур'є може мати багато процесів, які виконуються паралельно, в той час як програмне перетворення Фур'є, зазвичай, здійснює одну інструкцію за одиницю часу. Через це апаратні ШПФ-процесори можуть мати вищу пропускну здатність, ніж звичайні процесори. Паралельне виконання операції вимагає оптимального планування конвеєрів та часових параметрів, для того, щоб дані оброблялись вчасно для наступної стадії. Інші відмінності включають широке використання цілочисельної арифметики замість чисел з плаваючою комою і врахування доступних ресурсів FPGA-пристрою.

У роботі реалізовано 8-ми точкове ШПФ на FPGA-пристрої з використанням 12-бітних цілочисельних вхідних даних. 8-ми точкове перетворення було вибрано для того, щоб показати спосіб, який може бути використаний для генерування довгих ШПФ без візуалізації великих потоків даних. Нижче також будуть розглянуті результати з більших варіантів перетворень, реалізованих на FPGA-пристрої, та аналіз їх характеристик.

При реалізації швидкого алгоритму обробки сигналу на апаратній частині числа з плаваючою комою мають ряд недоліків:

- числа з плаваючою комою мають велику розрядність і займають багато комірок пам'яті;
- швидкість та затримка погіршуються через додаткову складність логіки;
- оскільки оцифровані сигнали мають фіксовану розрядність, то числа з плаваючою комою не мають переваг в обробці.

Тому тут використовуються 16-розрядні знакові числа з 12-бітними початковими даними, які мають діапазон значень від -1024 до 1024.

Привести уявну та комплексну компоненту експоненти до цілочисельного формату, описаного вище, можна наступним чином:

$$\begin{aligned} \text{real_part} &= \text{int}(\cos(2*\text{PI}*k/N)*1024), \\ \text{img_part} &= \text{int}(-\sin(2*\text{PI}*k/N)*1024). \end{aligned}$$

У модулі коефіцієнтів Фур'є (Twiddle factor ROM) (рис. 1) записані всі коефіцієнти Фур'є, які використовуються для перетворення. Кількість комірок (розмір) пам'яті та які значення в них записані залежать від розміру перетворення та типу перетворення.

У цій структурі двопортова оперативна пам'ять використовується, тому що базова операція ШПФ приймає два операнди на вхід та має два результати на виході, що дозволяє зчитувати операнди та записувати результати до пам'яті за один цикл. Також в структурі процесора ШПФ використовується два модулі двопортової пам'яті (2-port RAM), щоб зчитувати операнди першим модулем і одночасно записувати результати другим модулем оперативної пам'яті.

Модуль Butterfly Processing Unit (BFU) виконує ШПФ для двох точок (рис. 4), які вказані модулем генерування адрес.

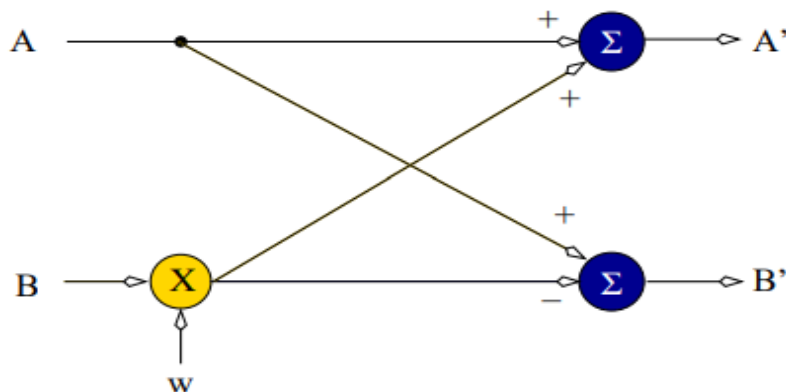


Рис. 4. Діаграма операції “метелика” в ШПФ.

Розрахунок виглядає наступним чином:

$$A' = A + B * W,$$

$$B' = A - B * W.$$

Для виконання цієї операції потрібно спочатку обчислити $B * W$, а потім здійснити додавання та віднімання. Формула цієї операції виглядає наступним чином:

$$B * W = (B_{real} + B_{img}) * (W_{real} + W_{img}) = B_{real} * W_{real} - B_{img} * W_{img} + B_{real} * W_{img} + B_{img} * W_{real}$$

Операція комплексного множення має два ступені: множення компонент та їх додавання між собою. Базові арифметичні операції (додавання, множення) зазвичай виконуються за один такт на FPGA-пристроях. Операції на одному рівні виконуються паралельно, що є перевагою перед програмною реалізацією. Таким чином, ця операція займатиме 2 такти. Крім того, на стороні операнда A буде потрібна затримка в два такти, щоб дійти до кінцевого обчислення вчасно разом з результатом множення.

Блок-схема даної операції на рівні передачі регістрів подана на рис. 5.

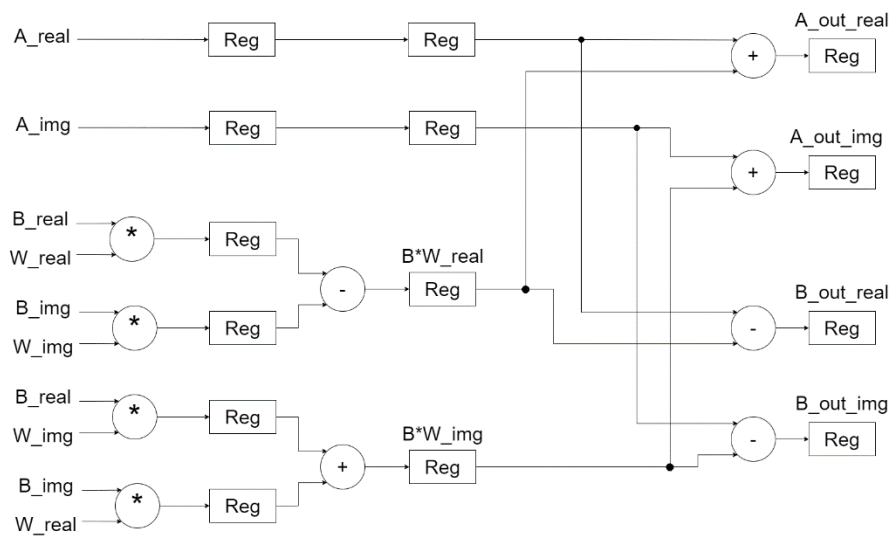


Рис. 5. Блок-діаграма структури модуля BFU.

Описана структура була реалізована на мові опису апаратних засобів VHDL в системі автоматизованого проектування Vivado IDE. Характеристики системи перевірялись на платі Nexys 4 DDR з FPGA-пристроєм сімейства Artix-7 фірми Xilinx [5]. Нижче наведено функціональну діаграму роботи модуля ШПФ.

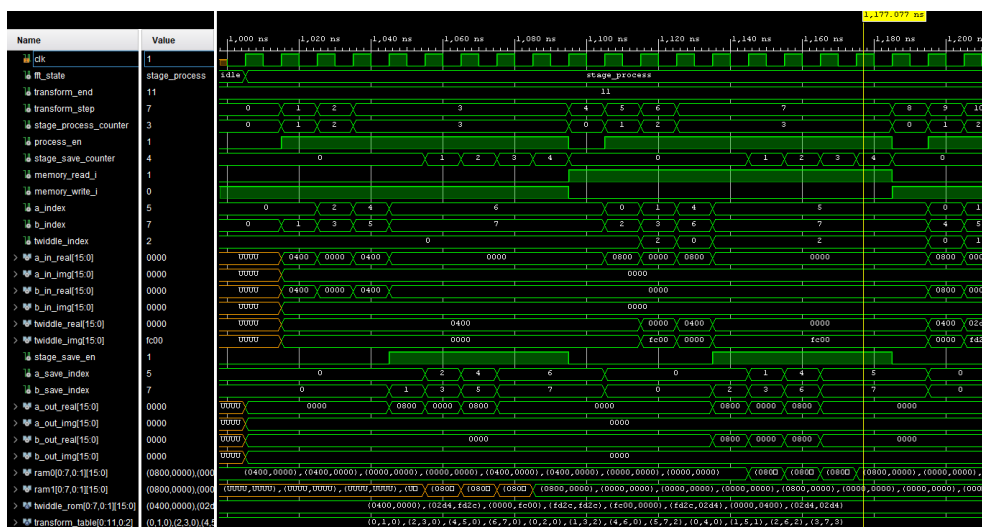


Рис. 6. Функціональна діаграма роботи модуля ШПФ.

На початку модуль переходить в стан обробки і здійснюється встановлення початкових сигналів. Спочатку на вхід модуля VFU надходять вхідні дані з першого блоку RAM через сигнали `a_in_real`, `a_in_img`, `b_in_real`, `b_in_img`. Через три такти з'являються відповідні вихідні дані та їх порядкові індекси для збереження. У той же час сигнал дозволу збереження (`stage_save_en`) стає "1", і починається запис вихідних даних в другий блок RAM-пам'яті. Коли всі дані були прочитані з першого блоку RAM-пам'яті, відбувається очікування на запис всіх вихідних даних в другий блок. Після того як всі дані були записані, сигнали `memory_read_i` та `memory_write_i` інвертуються і процес обробки стадії починається знову, але зчитування відбуватиметься з другого блоку пам'яті, а запис в перший блок.

Процес продовжується доки всі кроки перетворення не будуть виконані. У випадку 8-ми точкового перетворення кількість кроків перетворення 12 (від 0 до 11). Після цього модуль переходить в стан `done` (рис. 7).

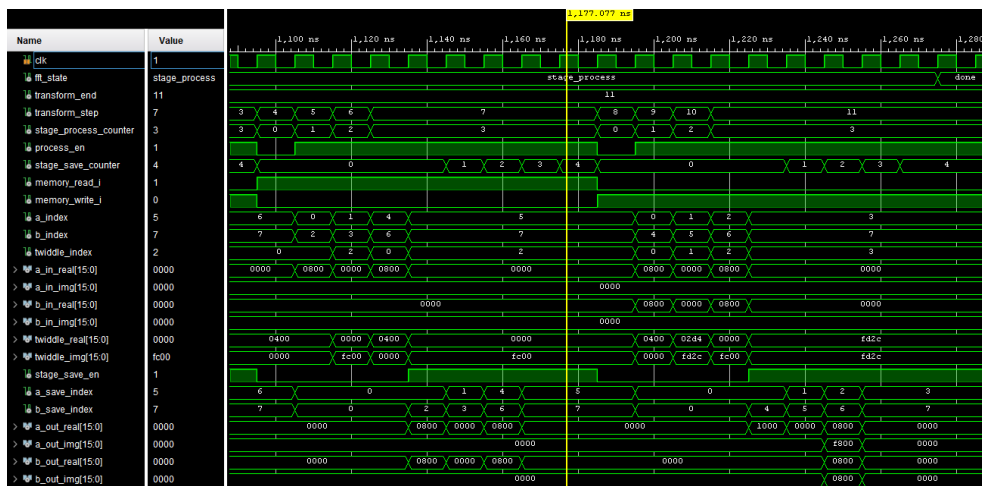


Рис. 7. Діаграма завершення роботи модуля ШПФ.

Для реалізації цієї структури на FPGA-пристрої необхідні 500 таблиць відповідності (LUT), 734 тригери, 4 DSP-блоки. У цьому випадку перетворення зайняло 270 нс при тактовій частоті 100 МГц.

Часові характеристики перетворень інших розмірів наведені в табл. 1.

Таблиця 1

Часові характеристики апаратного ШПФ різних розмірів при тактовій частоті 100 МГц

Розмір перетворення	Час перетворення
8	270 нс
16	520 нс
32	1 мкс
64	2,2 мкс
128	4,8 мкс
256	10,6 мкс
512	23 мкс
1024	51 мкс

Для прикладу, відомий мікроконтролер ESP32, який володіє стандартною тактовою частотою 240 МГц, виконує ШПФ розміром 1024 точок за 1,5 мс [6], що майже в 30 разів перевищує тривалість перетворення, реалізованого в роботі.

Таким чином, завдяки конвеєризації та паралельному виконанню команд апаратне ШПФ на FPGA є швидшим, ніж виконання на мікроконтролері або мікропроцесорі загального призначення. FPGA-пристрої дозволяють реалізовувати такі структури і роблять доступним швидкісну обробку сигналів та її застосування в системах реального часу.

IV. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Серед покращень, які можна реалізувати в перспективі, можна вказати заміну модуля ROM пам'яті, який зберігає експонентні коефіцієнти та збільшується із розміром перетворення, на модуль, який їх обчислює в реальному часі та має фіксований розмір. Одним із таких способів є використання алгоритму CORDIC [7], що дасть можливість зменшити кількість необхідної пам'яті для реалізації. Також можна розглянути схему з використанням лише одного блоку пам'яті, що призведе до зниження швидкодії перетворення, але і зменшить потреби оперативної пам'яті вдвічі.

V. ВИСНОВКИ

Проаналізовано спосіб реалізації швидкого перетворення Фур'є з використанням алгоритму Кулі-Тьюкі. Виконано апаратну реалізацію алгоритму швидкого перетворення Фур'є на мові VHDL та проведено синтез на базі ПЛІС сімейства Artix-7 (Xilinx) з використанням САПР Vivado IDE. Встановлено, що завдяки конвеєризації та паралельному виконанню команд реалізація апаратного швидкого перетворення Фур'є відбувається в 30 разів швидше, ніж його виконання на мікроконтролері чи мікропроцесорі загального призначення. Запропоновано варіанти для вдосконалення структурної схеми пристрою з метою підвищення швидкодії та зменшення пам'яті, необхідної для реалізації перетворення.

ЛІТЕРАТУРА

- [1] Lizhe Tan, Jean Jiang Digital Signal Processing. Fundamentals and Applications // Academic Press 2018, – 920 p.
- [2] K. Deergha Rao, M. Swamy Digital Signal Processing. Theory and Practice // Springer Singapore 2018, – 789 p.
- [3] G.W. Slade, "The Fast Fourier Transform in Hardware: A Tutorial Based on an FPGA Implementation", March 21, pp. 1-26, 2013
- [4] Ушенко Ю.О., Гавриляк М.С., Талах М.В., Дворжак В.В. Основи та методи цифрової обробки сигналів: від теорії до практики: навч. посібник. – Чернівці: Чернівецький національний університет імені Юрія Федьковича, 2021. 308 с.
- [5] Diligent, Inc. "Nexys 4 DDR Reference Manual." diligent.com. <https://diligent.com/reference/programmable-logic/nexys-4-ddr/reference-manual> (accessed Nov. 9, 2022).
- [6] Robin Scheibler. "FFT on the ESP32." robinscheibler.org. <https://www.robinscheibler.org/2017/12/12/esp32-fft.html> (accessed Nov. 9, 2022).
- [7] "CORDIC." en.wikipedia.org. <https://en.wikipedia.org/wiki/CORDIC> (accessed Nov. 9, 2022).

Automatic number plate recognition from videostream

Oleksandr Tsurkan, Yuliia Prokop

*Software Engineering Department
Odessa Polytechnic National University
Odesa, Ukraine*

Abstract — *Recognition of objects from images based on neural networks provide new opportunities in process automation. This paper considers the development of the server part of the service for recognizing car numbers from a video stream. Multithreading libraries, neural networks for image processing and video processing libraries are used for implementation.*

Keywords — *Computer vision, ANPR, Systems Analysis.*

I. INTRODUCTION

Nowadays IoT - internet of things - is gaining more and more popularity. With the help of a network of external cameras, it would be possible to create a system of accounting and control over the movement of cars. The cost of resources for processing video from many cameras is also important, so it will be about using only processors for inference of neural networks.

II. SCOPE

The aim of the work is to create an application for number plate recognition and analyze the speed of its execution on processors. An important aspect of the implementation is the speed of recognition on processors in order to reduce the cost of the necessary resources for execution. It is also important to recognize text from license plates of different lengths and different fonts, as license plates of different standards can be found on the roads. In the pipeline presented in this paper, the approach of dividing the task into subtasks is chosen, since accuracy plays a higher priority, a few fps are enough to recognize cars in the city.

III. RELATED WORK

There are two approaches to solve this problem. The first is an approach is end-to-end model, where one neural network, receiving an image as an input, outputs text and image coordinates as an output. The second approach is to divide the task into subtasks and use a separate neural network for each subtask. RpNet [1] is a end-to-end neural network, where network consists of two parts, first for number plate detection and second for character recognition. For the authors, the speed of the model was more important, what does this approach do better. The disadvantage of this approach is the lower quality of recognition. Efficient and layout-independent ALPR [2] is a model, which belongs to second approach. This model consists of the following steps: Vehicle Detection based on YOLOv2 [3], License Plate Detection and Layout Classification based on YOLOv2, and License Plate Recognition based on model that consists of the part of YOLO model. The second approach was chosen with dividing the task into subtasks. Accuracy plays a more important role, since a few frames per second are enough to recognize license plates in an urban area. We believe that car recognition is an unnecessary step in the approach given in the example of splitting the task into subtasks. YOLOv5 has enough generalization ability to distinguish a license plate from another random plate with text.

IV. METHODOLOGY

As input, we receive a video or video stream from the camera. To extract snapshots from a video will be used opencv2 library.

To recognize a car number from a snapshot, a pipeline from neural networks is used. First, it is required to recognize the segments of snapshots that supposedly contain license plates of cars. For this is used YOLOv5 model which is a Pytorch implementation of YOLOv4 [1]. The next step is to extract sections of the image segment that contain text. For this purpose was used CRAFT [2]. This is a neural network based on VGG16 [3], which has the main objective to localize individual character on image. For character recognition was used a neural network based on RESNET18 [4]. Residual networks use the approach of using outputs of a convolutional layer not only as inputs of the next one, but also as inputs of the layer located after the next one. To set up the system, you should test it on an existing device to determine how many frames to extract from one second of video. The test is to recognize license plates from 80 frames. Figure 1 shows a scheme for recognizing license plates with video.

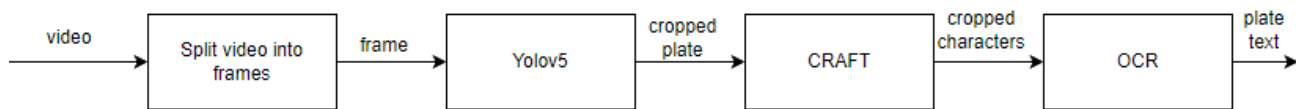


Figure 1. Scheme for recognizing license plates

V. RESULTS

The results of recognition from images are shown in figure 2. The time of recognition from one image is considered as the arithmetic mean of the prediction time for 40 images. One image is iteratively passed through the pipeline. Multithreading was tested on the threading library in python. The results are shown in figure 3.



Figure 2. Examples of numberplate recognition from some photos

Time for one frame reconition,s

Ryzen 5 1600 16 RAM DDR4	Ryzen 5 1600 16 RAM DDR4 2 threads	Xeon e5 2666 12 RAM DDR4	Xeon e5 2666 12 RAM DDR4 2 threads
0.33	0.31	0.5	0.38

Figure 3. Recognition time measurement results

VI. DISCUSSION

Based on the obtained results, you can get 2-3 frames per second. If you perform calculations on a high-end GPU, this result can be improved by about 10 times. The increased performance could be used for parallel processing of other videos, because increasing the FPS for a single video is not very practical.

VII. CONCLUSION

On tested devices, we got 2-3 fps, which is enough to recognize license plates of cars in the city if the camera has a top view over the road. One image is passed through the pipeline iteratively, if the batch approach is implemented, then the time for one snapshot is expected to decrease.

REFERENCES

- [1] Sai Sirisha Nadiminti, Pranav Kant Gaur, Abhilash Bhardwaj, "Exploration of an End-to-End Automatic Numberplate Recognition neural network for Indian datasets", arXiv:2207.06657v1, 2022
- [2] Rayson Laroca , Luiz A. Zanlorensi , Gabriel R. Gonçalves , Eduardo Todt , William Robson Schwartz , David Menotti1, "An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO detector", arXiv:1909.01754v4, 2021
- [3] Joseph Redmon, Ali Farhadi, "YOLO9000: Better, Faster, Stronger", arXiv:1612.08242v1, 2016
- [4] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection". arXiv:1512.03385v1, 2015
- [5] Youngmin Baek, Bado Lee, Dongyoon Han, Sangdoon Yun, Hwalsuk Lee, "Character Region Awareness for Text Detection", arXiv:1904.01941v1, 2019
- [6] Karen Simonyan , Andrew Zisserman , "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556v6, 2015
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition", arXiv:1512.03385v1, 2015

Using Goroutine to accelerate SQL queries for PostgreSQL in AWS

Olena Oleinik, Ihor Michurin

*Department of Software Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine*

Abstract—In the conducted study, the speed of execution of a series of SELECT SQL queries was compared when using the sequential format of code execution and parallel execution using Goroutine of the Golang programming language. As a result, it was established that the use of Goroutine reduces the time of execution of a series of requests and speeds up the performance of the code.

Keywords—*Software Engineering, Computer Science, Golang, Goroutine, Database management system, Software Tools, SQL, PostgreSQL, Amazon Web Services.*

I. INTRODUCTION

In the modern world, cloud services are actively used to host software systems and their components [1]. Nowadays, the most popular cloud services are Amazon Web Services, Microsoft Azure, and Google Cloud [2]. Parallel programming is also actively developing. The use of parallel programming tools is relevant both on servers and on personal computers [3]. Therefore, comparing the execution speed in serial and parallel formats for applications that interact with cloud services is essential.

II. RELATED WORK

The use of parallel programming tools in the Golang programming language is actively researched. Vincent Smith's "Go Web Scraping Quick Start Guide" [4] focuses heavily on the use of Goroutines and their advantages over sequential code execution. The book also shows potential problems that can arise when using parallel programming tools and shows ways to prevent such errors. In the work "Distributed computing in Go: comparative analysis of web application frameworks" [5] an evaluation and comparison of the characteristics of web frameworks of the Golang programming language was carried out based on defined criteria. A feature of the work is that a demonstration software system was created to perform the comparison of characteristics and the comparison was performed on its basis.

III. METHODOLOGY

To solve the problem, it was decided to compare the execution time of a series of SQL queries to the PostgreSQL DBMS [6] by sequential execution of queries and parallel execution of queries using Goroutine. The following environment was used for the research: the database is a cloud database of the db.t3.micro instance in the Amazon RDS service of the Amazon Web Services cloud with the PostgreSQL engine version 14.4 with 1 GB of RAM and 2 vCPU [7], the client is a personal computer with Golang programming language version go1.19.2 windows/amd64. During the study, 75 SELECT queries were executed sequentially against a PostgreSQL database table for sequential query execution. During the research, 75 Goroutines were run to execute queries in parallel, each executing a SELECT query against a PostgreSQL database table in parallel. The selected number of requests is due to the limit on the maximum number of simultaneous connections to the db.t3.micro instance.

IV. RESULTS

As a result of the research, the execution time of 75 SELECT SQL queries to the database in sequential format was 8.9189937 seconds, and the average execution time of one query was 118.221652 milliseconds. The execution time of 75 SELECT SQL queries to the database in a parallel format using the Goroutine tool was 775.031 milliseconds, and the average execution time of one query was 719.679689 milliseconds. It can be recommended to use Goroutine to execute a series of SELECT SQL queries to the database instead of sequential execution when using Golang.

V. DISCUSSION

To sum up, the total execution time of 75 SQL SELECT queries to the PostgreSQL database turned out to be faster during the parallel execution of queries using the Goroutine at 11.5 times. However, the average execution time of a single query during sequential execution was faster at 6.09 times. It is important to note that the number of completed requests is 75 which is relatively small. Therefore, an increase in the number of requests can lead to an increase in the accuracy of the conducted research. As a result of the study, a recommendation can be made to use parallel execution of SQL SELECT queries to speed up the overall speed of calculations.

VI. FUTURE RESEARCH

In further research, it will be advisable to increase the number of performed queries to the database to obtain more reliable results. Another direction is conducting research for various cloud services, such as Microsoft Azure, Google Cloud, and others. It will also be appropriate to execute the code in the Golang programming language on various clients, including cloud ones.

VII. CONCLUSION

Thus, it can be concluded that the parallel execution of SELECT queries using the Goroutine of the Golang programming language to the cloud database of the db.t3.micro instance in the Amazon RDS service with the PostgreSQL engine version 14.4 with 1 GB RAM and 2 vCPU speeds up the execution time of a collection of requests, but slows down the average execution time of a single request. It can be recommended to use Goroutine to execute a series of SELECT SQL queries.

VIII. ACKNOWLEDGMENT

The authors declare that there is no acknowledgment related to this paper.

IX. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] O. Alzakholi, L. Haji, H. Shukur, R. Zebari, S. Abas, and M. Sadeeq, "Comparison Among Cloud Technologies and Cloud Performance", *JASTT*, vol. 1, no. 2, pp. 40 - 47, Apr. 2020.
- [2] S. Kumar, H. Gautam, S. Singh, and M. Shafeeq, "Top vulnerabilities in cloud computing," *ECS Transactions*, vol. 107, no. 1, pp. 16887–16899, 2022.
- [3] A. Marowka, "Parallel computing on any desktop" *Communications of the ACM*, vol. 50, no. 9, pp. 74–78, Sep. 2007, doi: 10.1145/1284621.1284622.
- [4] V. Smith, *Go web scraping quick start guide: Implement the power of Go to scrape and crawl data from the web*. 2019.
- [5] E. Sochopoulos, "Distributed computing in Go: comparative analysis of web application frameworks," *dspace.lib.uom.gr*, 2020, Accessed: Nov. 18, 2022. [Online]. Available: <https://dspace.lib.uom.gr/handle/2159/24536>
- [6] The PostgreSQL Global Development Group, "PostgreSQL: The world's most advanced open source database," *Postgresql.org*, 2019. <https://www.postgresql.org/> (accessed Oct. 30, 2022).
- [7] "Amazon RDS Instance Types | Cloud Relational Database | Amazon Web Services," Amazon Web Services, Inc. https://aws.amazon.com/rds/instance-types/?nc1=h_ls (accessed Oct. 30, 2022).

Використання функціонального програмування у JavaScript та фреймворках

Олександр Жеребцов, Антон Іжиков

Кафедра інформаційних технологій
Запорізький інститут економіки та інформаційних технологій
Запоріжжя, Україна

Анотація—У наш час розробка веб-застосунків за допомогою фреймворків є актуальною задачею та має великий попит. Використання різних парадигм програмування у веб-розробці може якісно вплинути на процес створення даних застосунків. У роботі здійснено огляд механізмів та можливостей парадигми функціонального програмування (ФП) при веб-розробці у JavaScript, React та Next.js.

Ключові слова—Інженерія програмного забезпечення, Функціональне програмування, JavaScript, React.

I. ВСТУП

JavaScript – мультипарадигмова мова програмування. Це універсальна мова, що дозволяє використовувати кілька парадигм та програмувати у різних стилях. Наприклад, об'єктно-орієнтований стиль, де саме об'єкти, а не функції, є основним будівельним блоком, або процедурний стиль програмування, як в мові C, де функції є набором інструкцій, який може бути викликано з іншого місця в програмі. Недоліком такого мультипарадигмового підходу є те, що імперативний і об'єктно-орієнтований стиль припускають, що практично все в коді має бути змінним або схильним до мутації, мутабельним. Зазвичай об'єкти мають бути змінними, щоб значення їх властивостей могло бути змінено за допомогою методів.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

В імперативному програмуванні більшість структур даних є мутабельними, щоб забезпечити ефективну роботу з об'єктами та масивами [1]. Нижче наведено список речей, які притаманні деяким функціональним мовам і яких умовно немає в JavaScript:

- Чистота. У деяких функціональних мовах "чистота" підтримується самою мовою. Вирази із побічними ефектами неприпустимі. У JavaScript "чистота" може бути досягнута лише за згодою команди, тобто всі учасники повинні домовитися використовувати лише чисті функції. Якщо для більшої частини програми не використовується композиція чистих функцій, то це вважається недотриманням функціонального стилю.

- Незмінність (імутабельність). Мутація - це зміна значення структури даних без створення нової змінної та переприсвоєння значення. Приклад мутації у js наведено у рисунку 1.

```
helpers > JS example.js > ...  
55 const foo = { bar: 'baz' };  
56 foo.bar = 'qux'; // мутація
```

Рисунок 4.1 – Приклад мутації у js

Деякі ФП-мови блокують мутації. Замість мутацій існуючих значень структур даних, типу об'єктів і масивів, вирази поміщають результат у нові структури даних. Це може бути неефективно, але більшість ФП-мов усередині використовують спеціальні структури даних,

наприклад, префіксне дерево, з можливістю спільного використання даних. Тобто старий об'єкт і новий об'єкт зберігають посилання на ті самі дані, якщо вони не змінювалися. У ФП мовах імутабельність найчастіше дана за умовчанням. У JavaScript відсутні ефективні структури даних, які використовуються у більшості ФП-мов, але існують бібліотеки, які можуть допомогти у цьому питанні, наприклад `Immutable.js` та `Mori`. Є ознаки, що в майбутніх версіях специфікації ECMAScript очікується поява незмінних структур даних. Додавання `const` в ES6 є однією з таких ознак. Змінній, яка оголошена за допомогою `const`, може бути присвоєно інше значення. Важливо розуміти, що `const` робить значення змінної незмінним. Якщо змінна `const` вказує на об'єкт, значення змінної не може бути змінено повністю, проте можуть бути змінені властивості об'єкта. JavaScript має можливість заморозити об'єкт за допомогою `freeze()`, але такі об'єкти заморожуються тільки на верхньому рівні, а це означає, що якщо заморожений об'єкт має в будь-якій властивості інший об'єкт, то властивості цього іншого об'єкта піддаються змінам.

- Рекурсія – це здатність функції викликати саму себе. У багатьох функціональних мовах рекурсія – це єдина можливість виконати цикл. У таких мовах немає конструкцій на кшталт `for`, `while` чи `do...while`. Технічно JavaScript підтримує рекурсію, проте більшість функціональних мов мають таку особливість як "оптимізація хвостової рекурсії". Така особливість дозволяє рекурсивним функціям перевикористовувати фрейми стеку для наступних рекурсивних викликів. Фактично рекурсія перетворюється на плоску ітерацію. Без оптимізації хвостової рекурсії стек викликів зростає без обмежень і може спричинити переповнення стека. JavaScript з технічної точки зору має обмежену хвостову оптимізацію в стандарті ES6. На жаль, тільки один з найбільш поширених браузерів реалізував цю функціональність, а подібна оптимізація в Babel (найпопулярніший JavaScript компілятор, який використовується для компіляції ES6 в ES5), хоч і частково, була пізніше видалена. Тому наразі рекурсія все ще може бути небезпечною у використанні для великої кількості рекурсивних викликів.

З іншого боку, в JavaScript є:

- Функції першого класу.

Це використання функції як даних, тобто передавання функції як вхідних параметрів, повернення функції та присвоєння функції змінним та властивостям об'єктів. Ця властивість уможливує існування функцій вищого порядку, що, у свою чергу, уможливує появу часткового застосування, карування та композиції.

- Анонімні функції та лямбда-синтаксис. Наприклад, запис виду `<x=>x*2` є валідним виразом JavaScript. Такий синтаксис значно полегшує роботу з функціями вищого порядку.

- Замикання – це комбінація функції та її лексичного оточення. Замикання створюються під час створення функції. Коли функція створюється всередині іншої функції, то вона має доступ до змінних, оголошених у зовнішній функції, навіть після того, як буде здійснено повернення цієї зовнішньої функції. Замикання це те, що дозволяє працювати фіксованим аргументам часткових застосувань. Фіксований аргумент - це аргумент, заданий у контексті замикання функції, що повертається. Наприклад, у виразі `<add(1)(2)>` аргумент 2 є фіксованим аргументом для функції, що повертається під час виклику `add(1)`. Приклад замикання у js наведено у рисунку 2.

Мінливість об'єктів у JavaScript та сайд-ефекти можуть бути корисними. Сайд-ефект – це зміна зовнішнього стану, наприклад глобальної змінної, всередині функції. Створити сучасний додаток без сайд-ефектів наразі практично неможливо. Для вирішення цієї та вищеозначених проблем повністю функціональні мови на кшталт Haskell з коробки надають різні інструменти та математичні концепції, наприклад монади, функтори і т. д. JavaScript з коробки не дає такого різноманіття інструментів, але він має достатній набір ФП-властивостей, що дозволяє писати бібліотеки. І розбираючи ФП у JavaScript, неможливо не

загострити увагу хоча б на простіших поняттях типу монади та функтора. Але, як і більшість інших речей у функціональному програмуванні, академічні визначення зрозуміти складніше, ніж сприйняття ідеї. Наприклад, ознакою монади є ось це: "Монада в безлічі X це моноїд з категорії ендфункторів X , в якому морфізм, званий "твор", замінений композицією ендфункторів, а морфізм, званий "одиниця", замінений ендфунктором "тотожність" [2].

```
helpers > JS example.js > ...
57 const add = function( x ) {
58     return function( y ) {
59         return x + y;
60     };
61 };
62 // або
63 const addShort = ( x ) => ( y ) => x + y;
64 const summ = add( 1 )( 2 );
```

Рисунок 2 – Приклад замикання у js

Щоб зрозуміти, що саме це таке, потрібно зрозуміти, навіщо це потрібно. Відомо, що ФП передбачає обгортання функції практично всього поспіль. Доводиться писати багато маленьких функцій, що багаторазово використовуються, і викликати їх одну за одною, щоб отримати результат на кшталт (func1.func2.func3) або комбінації типу func1(func2(func3())). Та якщо все можна зробити шляхом комбінування набору функцій, то:

- Як обробляти умову if-else? (монада Either)
- Як обробляти винятки Null? (монада Maybe)
- Як переконатися, що функції, які дійсно багаторазово використовуються можуть використовуватися скрізь? (чисті функції (Pure functions), прозорість посилань (referential transparency)).
- Як переконатися, що дані, які ми надаємо у функції, не змінені і можуть використовуватися і в інших місцях? (чисті функції (Pure functions), незмінність)
- Якщо функція бере кілька значень, але у об'єднання в ланцюжок (chaining) можна передавати лише одне значення, то як зробити цю функцію частиною ланцюжка? (карірування та функції вищого порядку).

Задля вирішення цих питань в тому числі і існують монади. Плюс монади – це стандартизована річ, з простим інтерфейсом і передбачуваними властивостями. Архітектурні рішення в кожній компанії - свої унікальні, і наслідки від їх використання бувають різні. Також рішення залежать від обраної мови, фреймворку, тощо. Підсумовано, які зрештою існують основні тайпкласи і що вони вмюють, а також оглянемо деякі приклади. Для зручності скористаємось синтаксисом мови TypeScript, яка додасть типізацію динамічному js та допоможе розглядати інтерфейси як тайпклас.

Функтор – це базовий будівельний блок багатьох ФП термінів, в тому числі і монад. Це будь-який клас, що реалізує функцію Map певної сигнатури, на яку виконується одне просте правило. Тобто клас, або функція-конструктор, або тип даних, що зберігає значення та реалізує метод map. Наприклад, масив - це функтор, тому що він може зберігати значення і має метод map, що дозволяє нам застосувати (map) функцію до значень, що зберігаються. Такий клас дозволяє займатися мапінгом значення всередині контейнера, перетворюючи $T<A>$ в T. Наприклад, перетворення ітератора одних значень на ітератор інших значень; перетворення результату асинхронної операції [3].

Аплікативний функтор (Аплікатив) – це будь-який клас, що реалізує пару функцій Pure та LiftA2, для яких виконуються їх прості правила (в основному пов'язані з композицією).

Pure дозволяє створити контейнер, що містить єдине значення, LiftA2 дозволяє нам використовувати функцію від двох аргументів, маючи на руках два контейнери з відповідними типами, упакованими всередині. Реалізація цих методів гарантує автоматичну реалізацію тайпкласу "Functor". Цей клас дозволяє комбінувати разом пару незалежних обчислень $T<A>$ і T у загальний $T<C>$. Наприклад: зчеплення двох контейнерів (наприклад, List, Option, ZipList), парсинг мови із контекстно-незалежною граматику. "Аплікативний" означає, що з ним ми можемо застосовувати упаковані функції до упакованих значень. Наприклад, у нас може бути список функцій та список значень. Застосувавши до них LiftA2, ми отримаємо список результатів кожної функції, застосованої до кожного значення. Або виконати дві асинхронні операції та обчислити на їх підставі якусь відповідь.

Монада, якщо без «моноїдів у категоріях ендofункторів», – це будь-який клас з функціями Pure і Bind, яка приймає аргумент типу $T<A>$ і функцію, що перетворює розпаковане значення A в T, і повертає значення того ж типу T. Ідея сигнатури проста: є упаковане в контейнер значення типу A , є функція з голого A у такий самий контейнер, але вже зі значенням B , функція Bind дозволяє "зв'язати" ці два вирази разом, отримавши з пари $(T<A>, A \Rightarrow T)$ значення T. Реалізація цих методів гарантує автоматичну реалізацію тайпкласу "Applicative". Клас дозволяє комбінувати залежні обчислення, де T залежить від A , який у свою чергу знаходиться в контейнері того ж типу $T<A>$. Приклад - виконання кількох асинхронних операцій, залежних друг від друга; парсинг мови із контекстно-залежною граматику. Тобто, монади та інші тайпкласи - це прості інтерфейси, хай і досить абстрактні, які при цьому дуже потужні, і дозволяють зрозуміло висловлювати навіть складні речі.

III. МЕТОДИ

У даній роботі використовуються певні сучасні програмні комплекси для веб-розробки: React, Redux та Next.js.

React (старі назви: React.js, ReactJS) — відкрита JavaScript бібліотека для створення інтерфейсів користувача, яка покликана вирішувати проблеми часткового оновлення вмісту вебсторінки, з якими стикаються в розробці односторінкових застосунків. Розробляється Meta (раніше Facebook) і спільнотою індивідуальних розробників [4].

Redux — відкрита JavaScript бібліотека призначена для керування станом програм JavaScript[4]. Найчастіше використовується разом з React або Angular для побудови інтерфейсів користувача [5].

Next.js - відкритий JavaScript фреймворк, доповнення React.js для розробки веб-додатків, створений компанією Vercel (раніше ZEIT). Фреймворк був призначений для вирішення проблеми React.js, пов'язаної з відтворенням програми на стороні сервера - SSR. Працює на сервері та в браузері [6].

IV. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ

Як було показано, у поточному стані JavaScript досить непоганий для ФП. Netflix (і будь-яка програма на Angular 2+) використовує функціональні утиліти, засновані на бібліотеці RxJS. Facebook використовує концепт чистих функцій, функцій вищого порядку та компонентів вищого порядку для розробки Facebook та Instagram. PayPal, KhanAcademy і Flipkart використовують Redux для керування станом. Звісно, React не може бути «чистим» функціональним підходом. Але що від ФП він має?

- Функції можуть бути присвоєні іншим змінним, таким як інші типи даних, параметри або значення, що повертаються. Ця функція є у функцій js. Все, що підтримує js, підтримує

функціональне перетворення. Реалізація функцій зворотного виклику та замикань у js використовує цю можливість.

- Декларативне оголошення функції, а потім виклик функції для обчислення. Кожна частина має значення, що повертається, яке обчислюється безпосередньо.

Чиста функція без побічних ефектів. Тобто коли у функцію передаються однакові параметри, результат буде однаковим. Коли функція буде вводити той самий результат, а виводити різний? Коли зовнішня змінна використовується у функції, тому що зовнішня змінна може змінюватися наприкінці кожної запущеної програми. Отже, маємо універсальну функціональну формулу: не використовуйте, або хоча б не змінюйте зовнішні змінні всередині функції. Завдяки цьому функція може бути повністю незалежною ззовні, вільною від зовнішніх впливів і не торкатися зовнішнього середовища. Таким чином, виклик функції не матиме різних проблем, викликаних змінами у зовнішньому середовищі.

Багато методів, типу `reduce`, `map`, `find`, `forEach` та інші, підтримують специфікацію повернення нового об'єкту без зміни вихідного об'єкту.

У той же час в граматику ES6 було введено деякі нові граматики, які допомагають створювати і повертати новий об'єкт, наприклад оператор розповсюдження, шляхом деструктуризації та присвоєння новому об'єкту, який є відокремленим від вихідного об'єкта.

Компонент React – це генерація функції, а компонент вищого порядку в React – це функція вищого порядку. Наприклад, функції `connect` з бібліотек `Redux` та `React-router`. Фактично, це функція вищого порядку, яка розміщена поза додатком. Принцип той же, тобто дочірній компонент передається в батьківський компонент, а потім повертається новий компонент. Логіка батьківського компоненту може бути повторно використана, що саме собою також є своєрідною розв'язкою. Суть процесу рендерингу сторінки React – це вкладений процес виклику функцій.

Компонент передається батьківським компонентом крізь `props`. Це дає контроль, але субкомпоненти не можуть бути змінені безпосередньо всередині `props`, тобто єдиний потік даних, який у чомусь схожий на характеристики чистої функції та не змінює зовнішній стан. Тому, коли ми розробляємо компоненти, ми повинні намагатися щосили передати операції з побічними ефектами на зовнішній контроль, такі компоненти незалежні і легко адаптуються.

Функціональне програмування має функцію карування (`curry`), та виклик комбінації функцій (`compose`). Багато методів, які використовують `curry` та `compose`, часто зустрічаються в React та Redux, наприклад: `const middleware = store => next => action => {...}`. З цієї точки зору, React має характеристики функціонального програмування, а Redux - це модуль, що підключається, який суворо слідкує функціональному програмуванню.

Тож які плюси дає використання функціонального програмування у React?

- Швидкість розробки висока, а функції, що часто використовуються, можуть постійно повторно використовувати логіку. Функція є чорним ящиком ззовні і може використовуватися безпосередньо без побічних ефектів;

- Можна легко зрозуміти корисність через ім'я функції, не знаючи внутрішньої реалізації, наближеної до природної мови;

- Код більш зрозумілий, код, що фактично викликається, дуже простий, вся складна логіка інкапсульована в функції;

- Зручний для паралельної обробки, оскільки метод є чистою функцією і не впливає на зовнішні змінні, а послідовність обробки може бути організована довільно.

`Next.js` – це заснований на React фреймворк, призначений для розробки веб-застосунків, що мають функціонал, що виходить за рамки SPA, тобто так званих односторінкових

додатків. Як відомо, основним недоліком SPA є проблеми з індексацією сторінок таких програм пошуковими роботами, що негативно впливає на SEO. Втім останнім часом ситуація стала змінюватися на краще. Існують спеціальні інструменти, такі як react-snap, що дозволяють перетворити React-SPA на багатосторінник шляхом попереднього рендерингу програми на статичну розмітку. Метаінформацію можна вбудовувати в head за допомогою таких утиліт, як react-helmet. Однак Next.js істотно спрощує процес розробки багатосторінкових та гібридних додатків (останнього неможливо досягти за допомогою того ж react-snap). Він також надає багато інших цікавих можливостей. Таким чином, програми Next.js використовують усі переваги бібліотеки React і просто додають додаткові функції:

- Server Side Rendering. SSR дозволяє отримати доступ до всіх необхідних даних для побудови сторінки на сервері. Потім сторінка повністю відправляється назад у браузер і відразу відображається. SSR дозволяє веб-сторінкам завантажуватись за менший час і підвищує зручність роботи користувачів за рахунок підвищення швидкості відгуку;
- SEO або просто пошукова оптимізація. Використання SSR також надає перевагу у SEO, що допомагає сайту займати більш високі позиції на сторінках результатів пошукових систем. SSR підвищує рейтинг веб-сайтів для SEO, тому що вони завантажуються швидше та більше контенту сайту можна сканувати за допомогою трекерів SEO.
- Тег <head>. Next.js також дозволяє редагувати тег <head> сайту, чого неможливо зробити в React. Тег <head> є основною частиною метаданих веб-сторінки та сприяє підвищенню SEO-рейтингу сайту.

Основна перевага Next.js – вбудована підтримка SSR для підвищення продуктивності та SEO. Рендеринг на стороні сервера працює шляхом зміни потоку запитів React, так що всі компоненти, крім клієнта, відправляють свою інформацію на сервер. Маючи всю інформацію на сервері, клієнт може отримати попередню версію HTML-коду сторінки. Клієнт може надіслати один запит на сервер та отримати повноцінну HTML-сторінку замість того, щоб вимагати кожен компонент окремо з рендерингом на стороні клієнта.

V. ВИСНОВКИ

Починаючи з версії 7.1.0, React-Redux підтримує API хуків та надає такі хуки як useDispatch та useSelector, а починаючи з версії 5.1, з'явилася підтримка хуків і в React-Router. Екосистеми React та Next.js мають усі можливості для використання функціональної парадигми при створенні веб-застосунків. Розробники офіційно обіцяють і далі підтримувати класи, але відверто кажуть, що саме хуки стануть головним способом написання React-компонентів.

Насамперед, функціональний підхід загалом та хуки зокрема надають досить великі можливості. Потрібно добре розуміти, як вони працюють, особливо те, як вони мемоізують значення. Адже з великими можливостями приходить відповідальність. Якщо неправильно щось замемоізувати, то програма зберігатиме неактуальні значення, або навпаки — ререндер відбуватиметься надто часто. Щоб не зіштовхнутися з ситуацією, коли якесь значення замкнулося всередині чи у нижніх рівнях ієрархії, треба чітко розуміти, що таке замикання, як вони працюють, як функції обробляють вільні змінні тощо. Без цього знання буде посправжньому складно налагодити якусь проблему.

Та попри все, написання «функціонального» коду дає можливість поглянути на проблему з іншого боку, де розробка рішення може бути більш ефективною. І це просто збільшує кількість засобів висловити ідеї програміста. Відповідно, функціональне програмування може змінити стиль написання коду на краще. Якщо існуюча кодова база вже не дає необхідного рівня управління складністю бізнес-області, код засмічено непотрібними залишками, а імплементація нового функціоналу призводить до рефакторингу та втрати часу, перехід на функціональне програмування може бути рішенням. А в областях, пов'язаних з

великою кількістю обчислень або перетворень даних, паралельним чи асинхронним програмуванням, функціональне програмування напевно надає значних переваг.

ЛІТЕРАТУРА

- [1] В.В. Бублик, “Парадигми програмування” in *Об’єктно-орієнтоване програмування*, Київ, Україна: ІТ-книга, 2015, pp. 13-30
- [2] P. Henderson, *Functional Programming Application and Implementation*, NJ, USA: Prentice Hall, 1980
- [3] J. Hughes, *Why functional programming matters*, From “Research Topics in Functional Programming” ed. D. Turner, MA, USA: Addison-Wesley Longman Publishing Co, 1990, pp. 17-42
- [4] React, Wikipedia, 2022. Accessed on: Oct. 25, 2022 [Online]. Available: <https://uk.wikipedia.org/wiki/React>
- [5] Redux, Wikipedia, 2022. Accessed on: Oct. 25, 2022 [Online]. Available: <https://uk.wikipedia.org/wiki/Redux>
- [6] Next.js, Official website, 2022. Accessed on: Oct. 25, 2022 [Online]. Available: <https://nextjs.org/>

Контекст ІТ-проекту та проблеми бізнес-аналізу

Денис Гобов

*Кафедра інформатики та програмної інженерії
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського"
Київ, Україна*

Анотація— Ускладнення ІТ-проектів та поглиблення спеціалізації фахівців, що в них задіяні, призвели до розвитку інженерії вимог та бізнес-аналізу. Для вивчення поточного стану цих дисциплін в Україні проведено опитування фахівців українських ІТ компаній, що задіяні в процеси роботи з вимогами, та проведено статистичний аналіз залежності проблем проблем, з якими стикаються бізнес-аналітики та інженери з вимог, та контекстом проекту. Виявлено і досліджено статистично значущі залежності, що дозволяють передбачити виникнення певних проблем та забезпечити проактивні дії для їх послаблення або усунення.

Ключові слова—*Інженерія вимог програмного забезпечення, Бізнес-аналіз, Проблеми проектів, Контекст проектів.*

I. ВСТУП

Ускладнення проектів з побудови інформаційних систем та поглиблення спеціалізації фахівців призвели до формування таких дисциплін як «інженерія вимог» та «бізнес-аналіз». Їх вплив на успіх проектів підтверджено в наукових публікаціях з керування проектами. Наприклад, згідно з [1] найбільш впливовими чинниками, що забезпечують успіх проекту є залучення користувачів (15% респондентів), чітко сформульовані вимоги (13%), реалістичні сподівання щодо результатів проекту (8.2%), чітка постановка бачення та задач (2.9%). В рамках кожного проекту бізнес-аналітик має визначити підхід до бізнес-аналізу та спланувати свою роботу, враховуючи контекст проекту. Тому є важливим визначення актуальних проблем, з якими стикаються бізнес-аналітики, а також виявлення залежностей між проблемами та контекстом проекту. Розуміння цих залежностей дозволить передбачити найбільш вірогідні проблеми та знизити ризики їх появи завдяки проактивним діям як бізнес-аналітика, так і менеджера проекту. Дана робота присвячена вивченню досвіду практикуючих фахівців з України щодо проблем, пов'язаних з бізнес-аналізом, та впливу контексту проекту на них.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Більшість досліджень з даної тематики побудовані на базі аналізу результатів опитувань. В роботі [2] проведено опитування фахівців з Польщі, щодо виявлення найбільш поширених проблем щодо роботи з вимогами. В опитуванні взяли участь 55 бізнес-аналітиків, які обирали найбільш значимі проблеми із 64 запропонованих варіантів. До недоліків даного дослідження можна віднести недостатньо велику кількість респондентів, що унеможливило застосування методів статичного аналізу отриманих результатів. Серія досліджень була проведена в рамках міжнародної ініціативи NAPIRE, яка охопила більше ніж 10 країн. пов'язаних із вивченням проблем Ускладнення проектів з побудови інформаційних систем та поглиблення спеціалізації фахівців призвели до формування таких дисциплін як «інженерія вимог» та «бізнес-аналіз». В роботі [3] викладено результати опитування співробітників 228

компаній з різних країн світу щодо проблем, пов'язаних з інженерією вимог, та їх наслідків для ІТ-проектів. В роботі визначено 10 найпоширеніших проблем, а також здійснено причинно-наслідковий аналіз для них. Схожі дослідження проводились для Малайзії [4], Бразилії [5] та інших країн. Враховуючи, що Україна є одним з лідерів ІТ-ринку Європи та світу, є доцільним вивчення досвіду бізнес-аналітиків та інженерів з вимог цієї країни. Окрім того у наявних дослідженнях не застосовувались статистичні методи для оцінки впливу контексту проекту на проблеми.

III. МЕТОДИ

Для збору інформації щодо актуальних проблем, пов'язаних з бізнес-аналізом та інженерією вимог, було проведено опитування, в якому прийняли участь 328 фахівців з українських компаній, а також міжнародних компаній, що мають офіси розробки в Україні. Перелік проблем було сформовано із використанням робіт [2-5]. Учасники опитування мали обрати 5 найбільш критичних з їх точки зору проблем бізнес-аналізу, з якими вони зустрічались на проектах. Для визначення контексту проекту були використанні наступні характеристики:

- тип компанії: ІТ компанія (аутсорс, аутстафф, продуктова), не ІТ-компанія;
- розмір проекту;
- розмір компанії;
- клас систем або сервісів, що розробляються;
- розподіл команди (всі в одній локації або розподілена команда);
- досвід фахівця в бізнес-аналізі;
- методологія проекту (адаптивний чи предиктивний підхід);
- відповідальна за виявлення вимог роль (бізнес-аналітик/інженер з вимог, власник продукту, інша роль).

Структура, процес підготовки та перевірки опитувальника викладено в [6] та [7]. Для дослідження впливу контексту проекту на наявність проблем бізнес-аналізу було застосовано тест незалежності χ^2 -квадрат. Для цього було розроблено набір гіпотез про зв'язок між певним атрибутом проекту та наявністю певної проблеми.

- Нульова гіпотеза: не існує зв'язку між типом компанії, яка виконує проект, та наявністю проблеми «Відсутність практики верифікації вимог».
- Альтернативна гіпотеза: існує зв'язок між типом компанії, яка виконує проект, та наявністю проблеми «Відсутність практики верифікації вимог».

Будо підтверджено, що виконуються необхідні умови для застосування тесту χ^2 -квадрат:

- дані взяті випадковим чином із сукупності;
- розмір вибірки достатній – 324;
- змінні є взаємовиключними, жоден елемент не враховується двічі;
- значення вважаються адекватними, коли очікувана кількість не менше 5 і відсутні нульові значення.

Дослідження проводилось з використанням програмного забезпечення IBM SPSS Statistics 25.

IV. РЕЗУЛЬТАТИ

Згідно з результатами опитування найбільш поширеними є наступні проблеми: «Постійна зміна цілей проекту (процеси та вимоги)» - 43% респондентів, «Неповні або приховані вимоги» - 41%, «Недостатній доступ до інформації про потреби замовника та/або внутрішньої бізнес-інформації» - 33%, «Нестача часу», «Несумісні вимоги», «Формальна / недостатня валідація вимог з боку замовника», «Відсутність / ігнорування нефункціональних вимог» - по 27% кожна (Рис. 1). Найменш актуальними проблемами є «Недостатня підтримка лідером проекту» - 4%, «Проблеми термінології», «Технічно невиконувані вимоги», «Погані стосунки між замовником та менеджментом проекту» - по 8%.

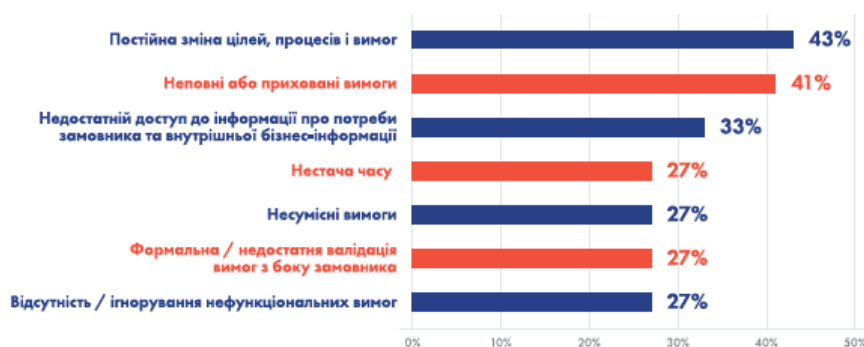


Рис 1. Найбільш поширені проблеми в бізнес-аналізі

Було виявлено ряд статично значущих залежностей між контекстом проекту та найбільш актуальними проблемами бізнес-аналізу (P-значення менше 0,05 з урахуванням рівня довіри 0,95):

- Тип компанії - «Недостатній доступ до інформації про потреби замовника та/або внутрішньої бізнес-інформації» (P=0,009);
- Розподіл команди - «Недостатній доступ до інформації про потреби замовника та/або внутрішньої бізнес-інформації» (P=0,003);
- Розподіл команди – «Нестача часу» (P=0,006);
- Категорія проекту - «Несумісні вимоги» (P=0,037);
- Розподіл команди – «Формальна / недостатня валідація вимог з боку замовника» (P=0,036);
- Розмір проекту – «Відсутність / ігнорування нефункціональних вимог» (P=0,03).

Для проблем «Постійна зміна цілей проекту (процеси та вимоги)» та «Неповні або приховані вимоги» статистично значущих залежностей не було виявлено.

V. ОБГОВОРЕННЯ

З урахуванням отриманих результатів по актуальності проблем та визначених залежностей між контекстом проекту та проблемами бізнес-аналізу, можна сформулювати наступні рекомендації:

- Бізнес-аналітики мають звертати уваги та розробляти заходи для зменшення ризиків виникнення проблем «Постійна зміна цілей проекту (процеси та вимоги)» та «Неповні або приховані вимоги»
- Більшу увагу ризику виникнення проблеми «Недостатній доступ до інформації про потреби замовника та/або внутрішньої бізнес-інформації» мають звертати бізнес-аналітики, що працюють в аутсорсінгових/аутстафінгових компаніях та розподілених командах.
- Більшу увагу ризику виникнення проблеми «Нестача часу» мають приділяти бізнес-аналітики, команда яких зібрана в одній локації.

- Ризик проблеми «Несумісні вимоги» менший для проектів розробки рішень з нуля, та найбільший для проектів по переробці інтерфейсу.
- Більшу увагу ризику виникнення проблеми «Формальна / недостатня валідація вимог з боку замовника» мають приділяти бізнес-аналітики, команда яких зібрана в одній локації.
- Більшу увагу ризику виникнення проблеми «Відсутність / ігнорування нефункціональних вимог» мають приділяти бізнес-аналітики, що працюють в команді із 15-30 фахівців.

Враховуючи динамічний розвиток ІТ-індустрії доцільним є регулярний перегляд виявлених залежностей з метою перевірки їх актуальності. На ризик виникнення також можуть впливати культурні особливості, тому результати дослідження не можуть бути екстрапольовані на інші країни.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальші дослідження можуть полягати у вивченні залежностей між використанням технік бізнес-аналізу, що визначені в [6], та проблемами бізнес-аналізу. Визначення цих залежностей дозволить побудувати рекомендаційну систему по уникненню проблем з врахуванням всього підходу до бізнес-аналізу на проекті.

VII. ВИСНОВКИ

Проведено дослідження стану бізнес-аналізу в Україні в частині проблем бізнес-аналізу в ІТ-проектах та визначено найбільш поширені проблеми. З використанням критерії хі-квадрат підтверджено гіпотезу щодо впливу контексту проекту на ймовірність виникнення проблем бізнес-аналізу. З врахуванням виявлених статично достовірних залежностей сформульовано рекомендації щодо можливих проблем бізнес-аналізу.

ЛІТЕРАТУРА

- [1] The Standish Group, *CHAOS Report*, The Standish Group, 2014.
- [2] A. Jarzębowicz, W. Ślesiński, “What Is Troubling IT Analysts? A Survey Report from Poland on Requirements-Related Problems,” *Advances in Intelligent Systems and Computing*, vol 830. Springer, Cham, pp. 3-19, 2019, doi: 10.1007/978-3-319-99617-2_1
- [3] D. Fernández, S. Wagner, M. Kalinowski, et al., “Naming the pain in requirements engineering: Contemporary Problems, Causes, and Effects in Practice”, *Empirical Software Engineering*, 22, pp. 2298–2338, 2017, doi:10.1007/s10664-016-9451-7
- [4] B. Solemon, S. Sahibuddin, A. Ghani, “Requirements engineering problems and practices in software companies: An industrial survey”, *International conference on advanced software engineering and its applications*, 2009, Springer, Berlin, Heidelberg, pp. 70-77.
- [5] D. Fernández, et al., “Naming the pain in requirements engineering: comparing practices in Brazil and Germany”, *IEEE Software*, 32 (5), PP.16-23, 2015, doi: 10.1109/MS.2015.122
- [6] D. Gobov, I. Huchenko, “Requirement Elicitation Techniques for Software Projects in Ukrainian IT: An Exploratory Study”, *Federated Conference on Computer Science and Information Systems*, 2020, pp. 673-681, doi: 10.15439/2020f16
- [7] D. Gobov, I. Huchenko, “Influence of the Software Development Project Context on the Requirements Elicitation Techniques Selection”, *Lecture Notes on Data Engineering and Communications Technologies*, vol 83, Springer, Cham, 2021, doi: 10.1007/978-3-030-80472-5_18

Створення багатофункціонального локального веб-серверу на базі ESP8266

Олександр Жеребцов, Марк Бороденко

*Кафедра інформаційних технологій
Запорізький інститут економіки та інформаційних технологій
Запоріжжя, Україна*

Анотація—Розглянуто питання використання вбудованих систем у якості багатофункціональних серверів малого навантаження. Обґрунтовано рішення розробити і запровадити клієнт-серверну систему з пристроєм моніторингу на базі мікроконтролера, розгорнутим на робочій станції, об'єднаними в локальну мережу. Описано та обґрунтовано вибір інструментів розробки для реалізації апаратно програмного комплексу, обрано апаратну частину проекту. Розроблено функціонал для взаємодії з сенсорами. Створено програмний код для управління мікроконтролером та програму для керування та обміном інформацією з мікроконтролерною системою.

Ключові слова—Комп'ютерна інженерія, веб-сервер, моніторинг даних, файловий менеджер, мультимедіа, чатування.

I. ВСТУП

У наш час використання Embedded systems та IoT є актуальною задачею, яка набуває стрімкого розвитку [1,2], а кількість «розумних приладів» для повсякденного приватного життя значно збільшується. Це досягається завдяки стрімкому прогресу у галузі комп'ютерних систем та відповідному здешевленню мікроконтролерів базового рівня та обладнання. Системи з моніторингу та контролю домашніх систем та систем невеликих підприємств чи офісів зможуть стати дійсними помічниками людини у його побуті та праці.

Однією з різновидів даних смарт-систем є «веб-сервер з моніторингу внутрішніх даних підприємства на базі ESP8266». Застосування розробки можливе у самих широких сферах автоматизації, наприклад, - системи «розумного дому» автоматизація тепличних господарств чи промислових об'єктів. Розробка може бути сумісна з роботами: «Мережа IoT на основі контролерів ESP8266» та «Система управління IoT мережею» [3,4].

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Повсюдне поширення бездротових мереж, розвиток хмарних обчислень, значна мініатюризація пристроїв послужили появі такого явища як Інтернет речей (Internet of Things, IoT). Це концепція обчислювальної мережі фізичних об'єктів («речей»), оснащених вбудованими технологіями для взаємодії один з одним або із зовнішнім середовищем, що розглядає організацію таких мереж як явище, здатне перебудувати економічні та суспільні процеси, що виключає з частини дій та операцій необхідність участі людини. Концепція сформульована у 1999 році як осмислення перспектив широкого застосування засобів радіочастотної ідентифікації для взаємодії фізичних об'єктів між собою та із зовнішнім оточенням. Наповнення концепції «інтернету речей» різноманітним технологічним змістом та впровадження практичних рішень для її реалізації почалося з 2010-х років і вважається висхідним трендом в інформаційних технологіях [1].

Вбудована система — це комбінація комп'ютерного обладнання та програмного забезпечення, призначеного для виконання певної функції. Вбудовані системи також можуть

функціонувати в рамках більшої системи. Системи можуть бути програмованими або мати фіксовану функціональність [2].

Вбудовані системи завжди функціонують як частина цілісного пристрою - це те, що мається на увазі під терміном *embedded*. Це недорогі, малоенергоспоживаючі маленькі комп'ютери, які вбудовані в інші механічні або електричні системи. Як правило, вони містять процесор, блок живлення, а також порти пам'яті та зв'язку. Вбудовані системи використовують комунікаційні порти для передачі даних між процесором і периферійними пристроями - часто інші вбудовані системи - за допомогою протоколу зв'язку. Процесор інтерпретує ці дані за допомогою мінімального програмного забезпечення, що зберігається в пам'яті. Програмне забезпечення зазвичай дуже специфічне для функції, яку виконує вбудована система [5].

ESP32 — це серія недорогих, малопотужних систем на мікроконтролерах з інтегрованим Wi-Fi і дворежимним Bluetooth. Серія ESP32 використовує мікропроцесор Tensilica Xtensa LX6 у двоядерних і одноядерних варіантах, двоядерний мікропроцесор Xtensa LX7 або одноядерний мікропроцесор RISC-V і включає вбудовані антенні перемикачі, радіочастотний балун, підсилювач потужності, малошумний підсилювач прийому, фільтри та модулі керування живленням. ESP32 створений і розроблений Espressif Systems, китайською компанією, що базується в Шанхаї, і виробляється TSMC, використовуючи їх 40 нм процес [6].

III. МЕТОДИ

У даній роботі було використано містить плату розробки ESP8266 Node MCU v.3 та датчик температури DS18B20.

Налагоджувальна плата Node MCU v.3 оснащена модулем ESP-12E, що містить мікросхему ESP8266 з мікропроцесором RISC Tensilica Xtensa® 32-bit LX106, який працює з регульованою тактовою частотою від 80 до 160 МГц і підтримує RTOS [7].

Також даний модуль має 128 КБ ОЗУ та 4 МБ флеш-пам'яті (для зберігання програм та даних), достатніх, щоб впоратися з великими рядками, які становлять веб-сторінки, даними в JSON/XML та всім, що ми сьогодні додаємо на пристрої IoT.

ESP8266 містить вбудований приймач Wi-Fi 802.11b/g/n HT40, тому він може не тільки підключатися до мережі Wi-Fi і взаємодіяти з інтернетом, але і встановлювати власну мережу, дозволяючи іншим пристроям підключатися безпосередньо до нього. Це робить ESP8266 NodeMCU ще більш універсальним.

Для програмування мікроконтролера було обрано середовище розробки Arduino IDE 1.8.19 з використанням бібліотек ESP8266WiFi.h [8], ESP8266WebServer.h [9], ESP8266SSDP.h [10], FS.h [11], OneWire.h [12].

Веб частину було розроблено за допомогою зв'язки HTML + CSS + JS у Visual Studio Code.

IV. РЕЗУЛЬТАТИ

Результатом розробки є веб-сервер, що складається з наступного функціоналу:

- головна сторінка (рисунок 1.);
- сторінка сенсорів та датчиків, які підключені до вбудованої системи;
- фотогалерея, з можливістю перегляду та завантаження зображень;
- система чатування всередині локальної мережі;
- файловий обмінник.

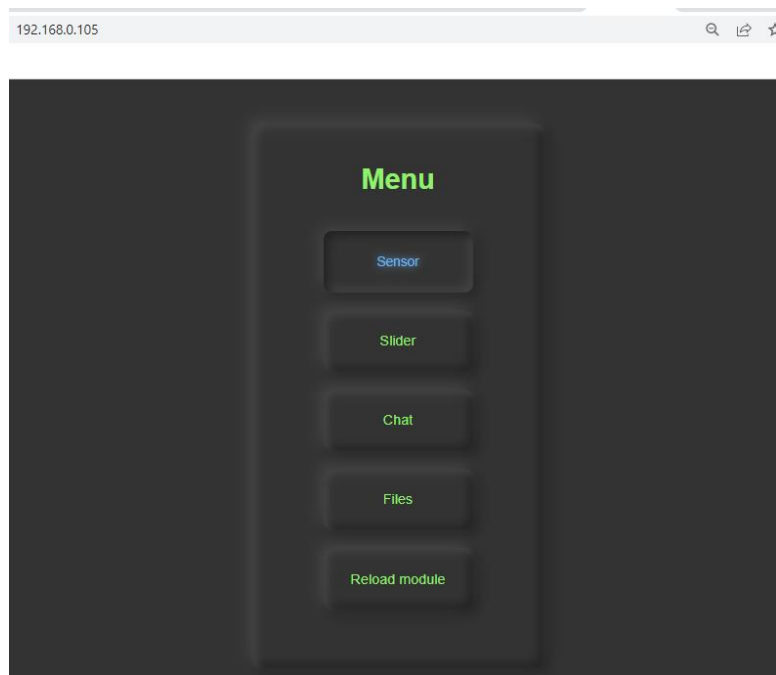


Рисунок 1 – Зовнішній вигляд головної сторінки

Користувач завдяки даним сторінкам може отримати інформацію про стан довкілля в осередку локальної мережі (у якості прикладу додано температурні показники), переглядати графічні зображення, чатуватись у середині мережі, використовувати внутрішнє сховище.

Даний пристрій може підключитись до WiFi-роутера існуючої мережі, або у разі її відсутності, організувати власну мережу, для підключення абонентів.

V. ОБГОВОРЕННЯ

Зовнішній вид сторінки сенсорів зображено на рисунку 2. Дані температури з датчиків можуть оновлюватись вручну (Refresh) користувачем та динамічно за часом. Route на отримання даних це "/data", далі спрацьовує функція обробник роуту "handleGetData", дана функція обробника формує JSON запит, який відправляється назад на клієнт. Усі запити відправляються видом "param", і в залежності від того дані з якого датчику треба отримати.

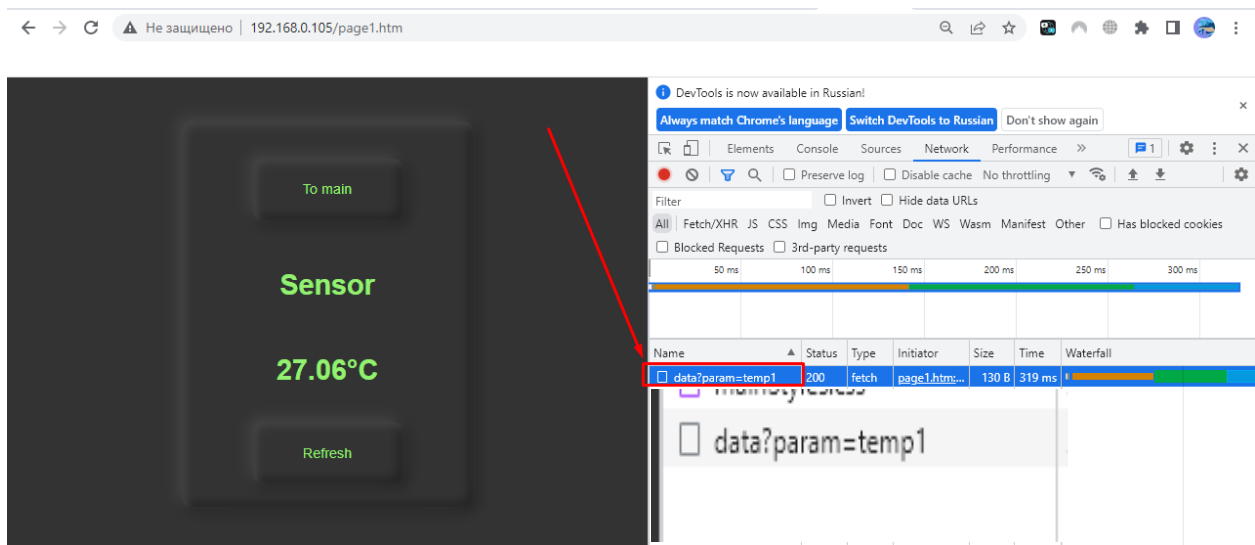


Рисунок 2 – Зовнішній вигляд сторінки сенсорів з демонстрацією передачі даних

Програмна частина слайдеру зображень наведена у лістингу 1, а результат роботи на рисунку 3.

Лістинг 1 – Програмна реалізація сторінки слайдеру

```
window.addEventListener('load', async () => {
  const carousel = document.querySelector('.carouselSlides')
  document.querySelector('.visibleContent').style.display = 'none';
  await getImages();
  document.querySelector('.visibleContent').style.display = 'block';
  document.querySelector('.loader').style.display = 'none';
  const leftBtn = document.querySelector('.leftBtn')
  const rightBtn = document.querySelector('.rightBtn')
  let images = document.querySelectorAll('.slide')
  let counter = 0;
  let step = images[counter].clientWidth / 2;
  const moveSlide = () => {
    carousel.classList.add('animation');
    carousel.style.transform = `translateX(-${step * counter}px)`;
    images.forEach(image => {
      image.classList.remove('activeSlide')});
    images[counter].classList.add('activeSlide');};
  async function uploadFile() {
    const imageInputElement = document.getElementById("image-file");
    const image = imageInputElement.files[0];
    if (!image) {alert('Select file first!')}return;}
    const formData = new FormData();
    formData.append("image", image);
    try {
      await fetch('/upload/image', { method: "POST", body: formData });
      alert('File downloaded!')
    } catch (error) {
      alert('File not loaded!')
    } finally {
      imageInputElement.value = ''
      imageInputElement.type = ''
      imageInputElement.type = 'file'}}
}
```

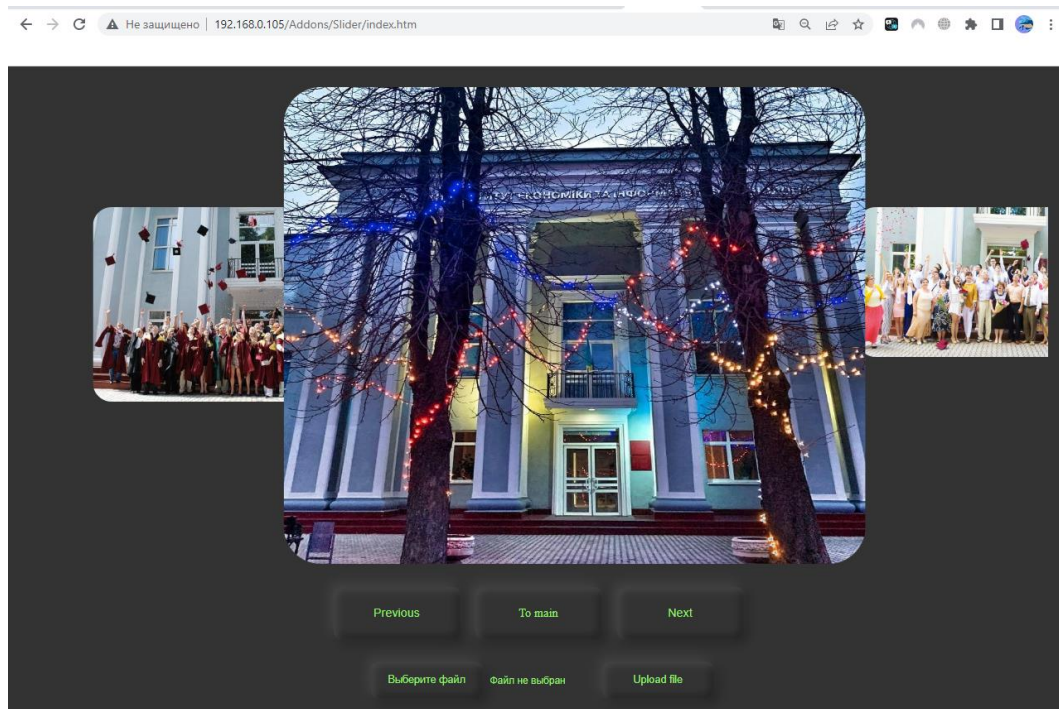


Рисунок 3 – Зовнішній вигляд сторінки слайдеру

Надалі особливу увагу слід приділити системі чатування (рис.4), частина коду якого наведена у лістингу 2.

Лістинг 2 – Програмна реалізація системи чатування

```
void HTTP_Data_Init(void){
  HTTP.on("/data",HTTP_GET, handleGetData);
  HTTP.on("/chat/messages",HTTP_GET, handleGetMessages);
  HTTP.on("/chat/message",HTTP_POST, handleAddMessage);
}, handleImageUpload);}
String messages = "";
void handleGetMessages(){
  String output = "{\"messages\":\":";
  output += messages;
  output += "\"}";
  HTTP.send(200, "text/json", output);}
void handleAddMessage(){
  if(HTTP.hasArg("plain")){
    messages+=HTTP.arg("plain");
    String output = "{\"messages\":\":";
    output += messages;
    output += "\"}";
    HTTP.send(200, "text/json", output);}}
```

Arduino HTTP сервер слідкує за GET запитами, які приходять з сторони браузера. Як тільки надходить повідомлення на адресу "/chat/messages", початок бере функція handleGetMessages(). Функція handleGetMessages() формує JSON строку, в яку передає всі повідомлення що є на даний момент, та відпраляє їх зворотно на клієнтську частину. Функція handleAddMessage() також формує JSON запит, куди збирає всі повідомлення, та ті що були дописані нещодавно. Та знову відпраляє їх на клієнтську частину.

Для запобігання перенавантаження NodeMCU, застосування JSON, було усунено, тому запроваджено для повідомлень зберігання у вигляді, який представлений у sendMessage.

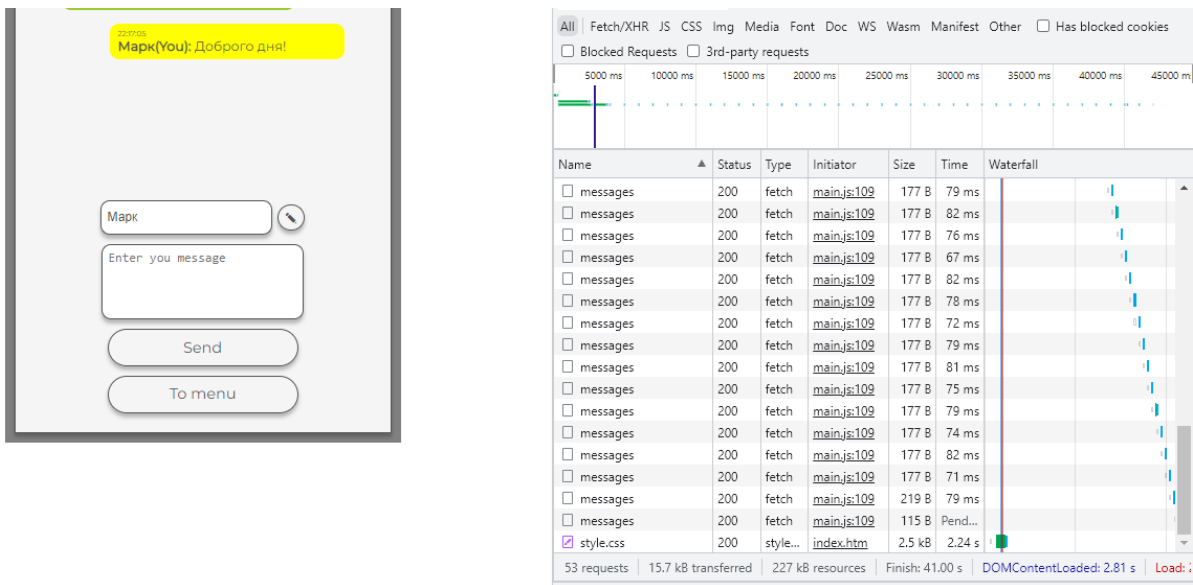


Рисунок 4 - Зовнішній вигляд сторінки чатування та демонстрація обміну повідомленнями

Сторінка «Files» (рис. 5), є бібліотекою, за допомогою якої можливе редагування, видалення та скачування файлів, відповідальних до веб інтерфейсу. Було вирішено використати дану бібліотеку, за для зручності та підвищенню функцій веб-вузла.


```
1 <!doctype html>
2 <html lang="ru">
3 <head>
4 <meta charset="utf-8" name="viewport" content="width=device-width, initial-scale=1.0, user-scalable=no" />
5 <title>Главная</title>
6 <link rel="stylesheet" href="mainStyles.css" />
7 </head>
8 <body>
9
10 <div class="menuContainer">
11 <h1>Главная</h1>
12 <a class="noselect blue" href="/page1.htm">Датчик 1</a>
13 <a class="noselect blue" href="/Addons/Slider/index.htm">Слайдер</a>
14 <a class="noselect blue" href="/Addons/Chat/index.htm">Чат</a>
15 <a class="noselect blue" href="/edit">Редактировать</a>
16 <button class="noselect red reFBtn" onclick="refreshModule()">Перезагрузить модуль</button>
17 </div>
18 <script>
19 function refreshModule(){
20   if (confirm("Вы уверены?")) {
21     fetch("/restart?" + new URLSearchParams({
22       device: 'ok',
23     }
24     ))
25     .then(res=>window.location.reload());
26   }
27 }
28 </script>
29 </body>
</html>
```

Рисунок 5 - Зовнішній вигляд сторінки роботи з файловою системою

VI. ВИСНОВКИ

У результаті було здійснено проектування апаратної та програмної частини вбудованої системи з моніторингу даних локальної мережі. Було створено веб-сторінку, для взаємодії з серверною частиною. Було розроблено метод обробки датчиків температури, що дозволяє опитувати їх без використання функції затримки у програмному коді мікроконтролера. Розроблено систему моніторингу мультимедійного контенту та систему яка дозволяє відправляти повідомлення. Проект буде корисним для невеликих офісних компаній у якості дашборду роботи з проектами.

ЛІТЕРАТУРА

- [1] Antar Shaddad Abdul-Qawy, P. J. Pramod, E. Magesh, T. Srinivasulu, "The Internet of Things (IoT): An Overview," *Int. Journal of Engineering Research and Applications*, vol. 5, issue 12, (part - 2), Dec., pp.71-82, 2015
- [2] F. Samie, L. Bauer, J. Henkel, "IoT technologies for embedded computing: A survey," *Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES '16)*, 2016, pp. 1-10, doi: 10.1145/2968456.2974004
- [3] А. В. Сагун, В. В. Хайдуров, І. А. Поліщук, "Розробка програмних модулів для обміну даними у промислових мережах," Київ, Україна: НТУУ «КПІ» ім. Ігоря Сікорського, 2021. Accessed: Oct. 10, 2022. [Online]. Available https://ela.kpi.ua/bitstream/123456789/46647/1/RPMODPM_posibnyk.pdf.
- [4] С. С. Шлапак, "Система управління IoT мережею," магістерська дисертація, Факультет інформатики та обчислювальної техніки, НТУУ «КПІ» ім. Ігоря Сікорського, Київ, Україна, 2019.
- [5] E. White, "Making Embedded Systems:" *Design Patterns for Great Software*, CA, USA: O'Reilly Media, 2011
- [6] ESP32, Wikipedia, 2022. Accessed on: Oct. 10, 2022 [Online]. Available: <https://uk.wikipedia.org/wiki/ESP32>
- [7] NodeMCU V3 ESP8266 12e Development Board CH340 (2022). Accessed on: Oct. 10, 2022. [Online]. Available: <https://circuit.rocks/nodemcu-v3-esp8266-development-board-ch340.html>
- [8] Arduino core for ESP8266 WiFi chip, GitHub, 2022. Accessed on: Oct. 10, 2022 [Online]. Available: <https://github.com/esp8266/Arduino>
- [9] ESP8266WebServer, GitHub, 2022, Accessed on: Oct. 10, 2022 Available: <https://github.com/esp8266/Arduino/tree/master/libraries/ESP8266WebServer>
- [10] ESP8266SSDP, GitHub, 2022, Accessed on: Oct. 10, 2022 [Online]. Available: <https://github.com/esp8266/Arduino/tree/master/libraries/ESP8266SSDP>.
- [11] File system wrap, 2022. [Online]. Available: <https://github.com/esp8266/Arduino/blob/master/cores/esp8266/FS.h>. Accessed on: October 10, 2022
- [12] OneWire, Arduino – Home, 2022. Accessed on: October 10, 2022. [Online]. Available: <https://www.arduino.cc/reference/en/libraries/onewire/>

Selection of Defective Chains in a PCB Image by Thinning, Flood-Filling and Overlaying

Roman Melnyk, Ruslan Tushnytskyy, Andriy Shpek

*Software Department
Lviv Polytechnic National University
Lviv, Ukraine*

Abstract—The flood-fill algorithm is used to color and to separates the PCB chains components. The thinning algorithm to find specific points to be starting positions in the flood filling algorithm is realized.

Keywords—*Printed Circuit Board, Defect Detection, Chain, Short, Flood-Fill, Thinning.*

I. INTRODUCTION

The main part of defects in manufactured PCB samples relates to their chains realized by traces and contacts. Defects are of different nature. One group of them relates to extra conducting material or a lack of conducting material. The second group unites defects caused by incorrect shape of traces and different sizes which are also caused by a lack or extra amount of conducting material. In PCB images all these defects are reflected by black and white pixels arising in comparison to a template sample.

Today, there are specialized optical systems that make it possible to control the quality of manufactured boards. However, one cheaper option is to use an image comparison of a reference image of the board with a manufactured one. During the comparison of these images, the tasks of detecting the location with a defect, evaluating the interpretation of the defect, evaluating the accuracy of the defect detection, and illustrating the location of the defect detection on the board arise.

II. RELATED WORKS

Some works considering defects in PCB images and method of their extraction are given in the survey [1-2]. In [3] the approach of PCB different defects extraction are considered and divided in groups of image subtraction and features extraction. The Image subtraction methods are the simplest approaches used for the PCB images control. The manufactured PCB images are compared with the template image and the difference image contains positions with defects. Algorithm in [4] to increase visibility of defects uses parts of the board. Simple techniques to divide a plane are used. In work [5] authors propose an algorithm for machine vision based PCB defect inspection system to detect the defects viz. short circuit, open circuit and depressions only. The work [6] considers parallelization issues for intrinsically sequential algorithms of thinning. Algorithms of thinning, clustering and comparison are used in the work [7] to find defective contacts and traces.

In this article three approaches are considered partition by multilevel segmentation of cumulative histogram, flood filling to select and separate chains and measurement of traces to detect positions and intensity of PCB defects.

III. Methodology

An example of the PCB image with two defects is shown in Fig. 1(a). Open and short defects are encircled by a red color. In the first case a small area of a black trace is changed by white pixels. And vice versa white pixels of the space between the chains changed their color to black.

As a rule, several pixels reflecting defects is very small comparatively to a few white or black pixels of the whole image. Even having a template sample, they are superimposed on the inaccuracies of manufacturing PCBs. Also, they are small comparatively to a size of a row or column in the image matrix. That is why the main problem to be solved is a selection of defective chains from the whole PCB image to decrease influences of all surrounding chains. The preliminary step of the algorithm to solve this problem is a selection of all chains in the template PCB image. The selection and separation of chains are realized by the flood filling algorithm. Arbitrary black pixels in Fig.1(a) are accepted as start points for the algorithm. The resulting PCB image is shown in Fig. 1(b).

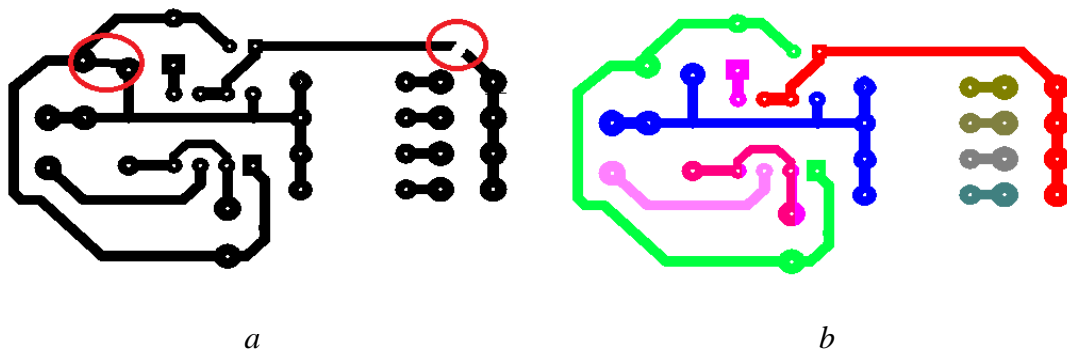


Fig. 1. Image of a PCB with two defects (a); the etalon PCB image after flood filling (b)

To find identification points belonging to separated chains the thinning algorithm is applied to the colored PCB image. Then coordinates of specific points of every chain are found. For the considered case they are as switches i.e., points in which three and more lines are merged. The skeleton image is colored too. It is illustrated in Fig. 2(a) by an overlay of the colored PCB image and the skeleton image.

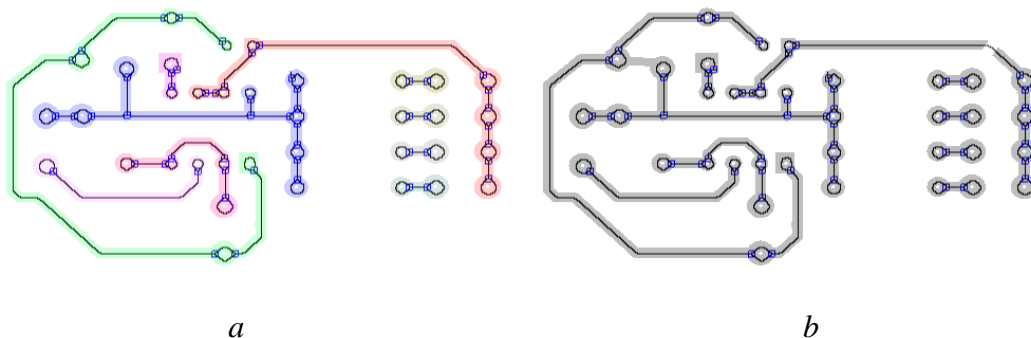


Fig. 2. The overlay of the etalon and skeleton PCB images (a); the overlay of the defective and skeleton PCB images (b)

Then the skeleton image is overlaid with the defective PCB image with black chains. The result image is shown in Fig. 2(b). This overlaying allows to project specific points of the etalon image on corresponding pixels of manufactured PCB image.

Thus, two sets of pixels are formed:

1. $E_e = \{E_1, E_2, \dots, E_n\}$ is a set of specific points for the etalon PCB image and E_i is a set for one chain.
2. $D_e = \{D_1, D_2, \dots, D_n\}$ is a set of specific points for the manufactured PCB image and D_i is a set for one chain.

IV. RESULTS

Now, by choosing start points from the D_i set it is possible to select and separate the chain containing pixels belonging to D_i . The flood filling algorithm takes one point position from D_i and changes a color of the chain of the manufactured PCB image. Depending on selected start points two cases are possible. The chain having an open defect can be colored to the defect border. These two colored parts are shown in Fig. 3 (a, b). In both cases for different defects selected and separated defective chains are input data for the following algorithms comparing template and defective chains of the PCB image.

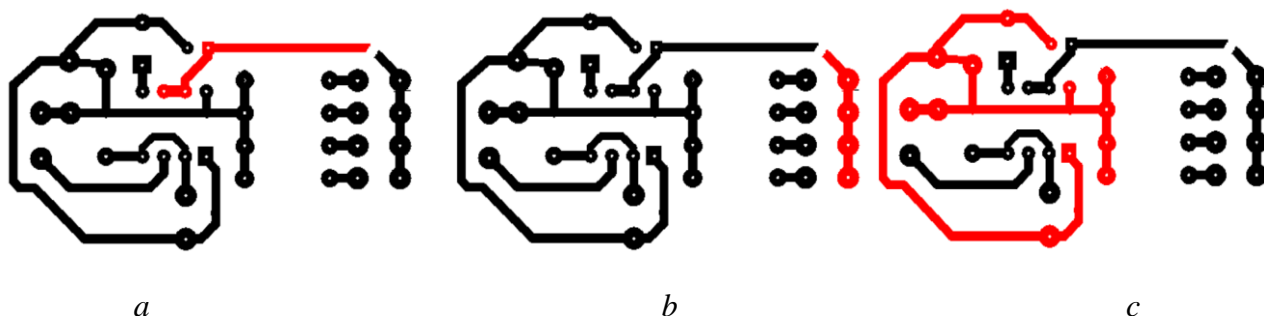


Fig.3. Two cases of colored defective chain (a, b); colored chains with defects (c)

For the short defect in the PCB image the flood filling algorithm colors a component containing two or more chains. This case is illustrated by colored connected two chains in Fig. 3 (c). The following algorithms comparing template and defective chains of the PCB image.

V. DISCUSSION

Let's discuss evaluation methods for detecting defects: histogram & gradient measurements.

Histogram measurement. Flood filled by different colors chains from the etalon and real PCBs are objects for comparison. Comparison can be realized for all components or only to those not satisfying checking conditions. There are many approaches allowing to measure and to compare robust and suspicious components. Some of them are presented here.

To illustrate the main idea an ordinary special histograms are calculated in which instead of traditional probability/intensity axes probability/coordinate axes (OX or OY) are used:

$$G_c(s) = \sum_{j=1}^s G(j), \quad j=1, \dots, s, \quad s=1, \dots, N \quad (1)$$

where $G(j)$ is the overall number of pixels in PCB elements for the j -th coordinate, $G(s)$ is a cumulative number of pixels for the s -th coordinate, N is the number of intervals.

In Fig.4 two chains are shown. The first one is selected from the etalon PCB image and the second has defects within its trace.

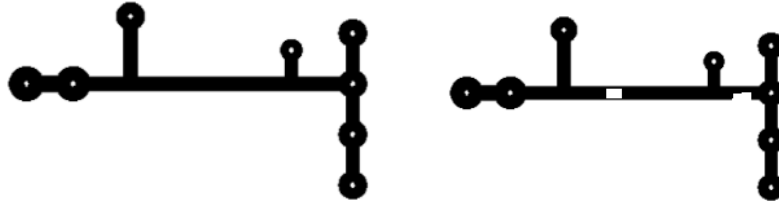


Fig. 4. Two chains of a PCB images

In Fig. 5 two special ordinary histograms for chains from the PCB image by OX axis are shown. In the defective intervals graphs differ between themselves by 20 percent. So, defects can be easily detected and measured.

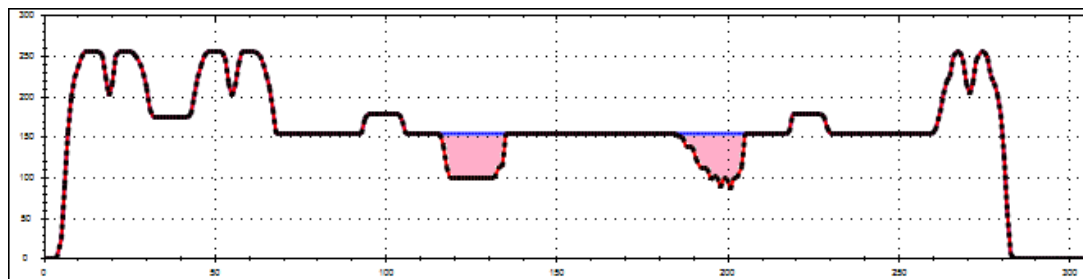


Fig. 5. Ordinary histograms of pixels in traces by OX axis

Additionally in Fig. 6 two special ordinary histograms for chains from the PCB image by OY axis are shown. In the defective intervals graphs differ between themselves by 10 percent. So, these graphs also confirm defects and determine their coordinates.

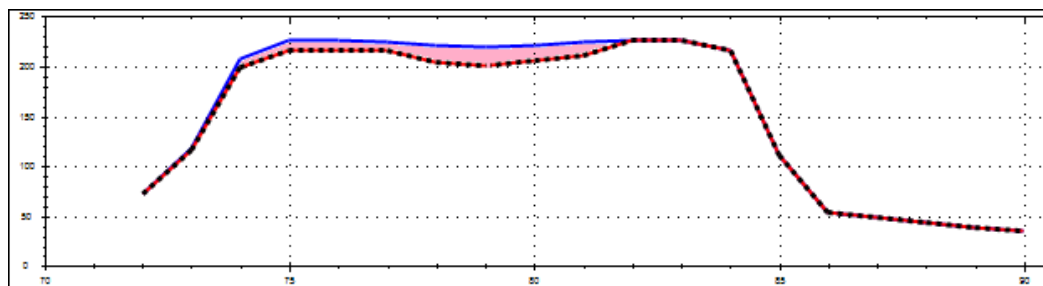


Fig. 6. Ordinary histograms of pixels in traces by OY axis

Gradient measurement. One more function to estimate defects is to measure mean gradient in rows or columns:

$$\bar{I}(i) = 1/W(\sum_{j=1}^W I(i,j)), i = 1,2,\dots,H, \quad (2)$$

$$G(j) = 1/H(\sum_{i=1}^H G(i,j)), j = 1,2,\dots,W, \quad (3)$$

where $G(i, j)$ is gradient value for two neighbouring pixels in the i -th row and the j -th column ($1 \leq i \leq H, 1 \leq j \leq W$).

For two chains their mean gradient functions are given in Fig.8.

The red graph of the defective chain has some deviation in places of defects in a body of the chain. In Fig.7 absolute values of the template and defective image are 3.8 and 6.5 for corresponding coordinates. It is relative deviation near 70 %. For not inverted images these values

are the same. So, in the case of gradients there is no difference what types of images are being processed.

Such a procedure is being applied for every chain which the connectivity algorithm marks as suspicious. To make a full guaranty it can be applied to all chains of the PCB image.

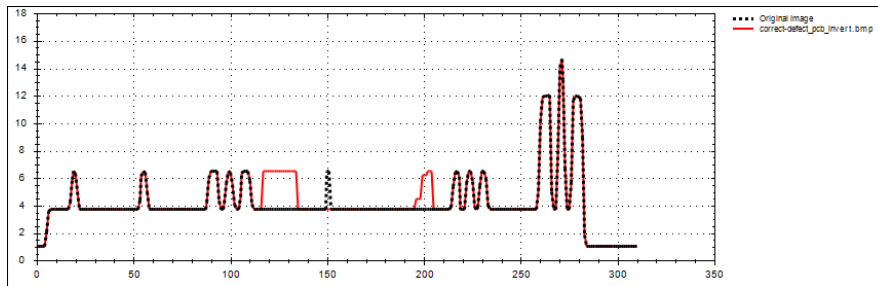


Fig. 7. Mean gradient functions for template chain and defective chain

To measure the difference between graphs of the real and etalon components the distance D_{ch} between two graphs is calculated as the sum of an absolute (or quadrate) value of difference between the functions for every intensity value:

$$D_{ch} = \sum_{s=1}^W |f_1(s) - f_2(s)| \quad (4)$$

Having values of the manufacture tolerance and calculated D_{ch} the program takes a decision: to reject or to accept the current PCB fragment as suitable.

VI. FUTURE RESEARCH

Further work includes optimizing the work of the developed algorithms, finding new parameters for effective and high-quality defect detection.

VII. CONCLUSION

The flood-fill algorithm is realized and applied to the PCB image to separate one connection wire from it. The thinning algorithm is realized to find specific points to be starting positions in the flood filling algorithm. An approach based on overlaying images is used to project specific points on defective chains. It helps to indicate the wire breaks, short circuits, and others type of defects.

REFERENCES

- [1] M. Moganti, F. Ercal, C. H. Dagli, S. Tzumeckawa, "Automatic PCB inspection algorithms: A review", Computer Vision and Image Understanding, vol.63, issue 2, pp.287-313, April 1996.
- [2] D.B. Anitha, M.Rao, "A survey on Defect Detection in Bare PCB and Assembled PCB using Image Processing Techniques", Int. Conf. on Wireless Communications, Signal Processing and Networking, pp.39-43, 2017.
- [3] J. Nayaka, K. Anitha, B.D. Parameshachari, R. Banud, P. Rashmi, "PCB Fault Detection Using Image Processing", IOP Conf. Series: Materials Science and Engineering, vol. 225, pp. 1-5, 2017.
- [4] Y. Hanlin, W. Jun, "Automatic Detection Method of Circuit Boards Defect Based on Partition Enhanced Matching", Information Technology Journal, vol.12(11), pp. 2256-2260, 2013.
- [5] S. Guan, F. Guo, "A New Image Enhancement Algorithmfor PCB Defect Detection", Intern. Conf. Intelligence Science and Information Engi-neering (ISIE), pp. 454-45, 2011.
- [6] F.B. Nadaf, V.S. Kolkure, "Detection of Bare PCB Defects by using Morphology Technique", Intern. Journal of Electronics and Communication Engineering, vol. 9, n. 1, pp. 63-76, 2016.
- [7] R. Melnyk, D. Hatsosh, Y. Levus, "Contacts detection in PCB image by thinning, clustering and flood-filling", IEEE 16th International conference CSIT 2021, pp. 370-374, 2021.

Автоматизоване тестування в Selenium: Page Object та Page Factory

Віра Гарасимів

*Кафедра комп'ютерних систем і мереж
Івано-Франківський національний технічний університет нафти і газу
Івано-Франківськ, Україна*

Анотація—Основною метою даного дослідження є підкреслення переваг використання патерну Page Object для написання авто-тестів. Кожна окрема сторінка веб-додатку представляється окремим класом, в якому зберігаються елементи, що знаходяться на цій сторінці та методи, що будуть працювати з ними. Для створення та ініціалізації об'єктів, що реалізують патерн проєктування Page Object, використано Page Factory.

Ключові слова—Патерн проєктування, Page Object, Page Factory, Веб-додаток

I. ВСТУП

Написання автоматизованих тестів – це необхідність і важливий інструмент для швидкого пошуку дефектів на початкових етапах розробки програмного забезпечення. Коли з'являється нова функція на ранньому етапі розробки, розробники можуть запустити автоматизовані тести і перевірити, чи нові зміни вплинуть на інші компоненти програмного забезпечення.

Хоча написання автоматизованих тестів може здатись досить легкою задачею, проте існує можливість одержати погано реалізовані тести і високу вартість їхньої підтримки. Тому широкого використання набули моделі Page Object для написання підтримувальних та багатократно використовуваних тестових процедур в Selenium.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Задач в області автоматизації проєктування і розробки є багато, і зустрінеш з ними, програмісти сформували патерни. Спочатку класичні патерни були сформовані в роботі «Design Patterns: Elements of Reusable Object-Oriented Software», де були описані усі патерни, які могли вирішити поставлені завдання [1]. Основна проблема патернів заключається в їхньому коректному використанні, необхідно розуміти, що використання патерну повинно, в першу чергу, вирішити поставлене завдання. Патернів існує дуже велика кількість. Наприклад, в книзі [2] їхня кількість сягає 86. Найпоширеніші патерни автоматизованого тестування описано у праці [3].

III. МЕТОДИ

Page Object – це патерн проєктування об'єктів в Selenium, у якому веб-сторінки представлені у вигляді класів, а різноманітні елементи на сторінці визначаються як змінні відповідного класу. Оскільки ми використовуємо мову програмування java, то створюємо Maven проєкт в IntelliJ Idea та просимуємо залежності в POM-файл. Для створення тестового сценарію використано фреймворк TestNG:

```
<dependency>
  <groupId>org.seleniumhq.selenium</groupId>
  <artifactId>selenium-java</artifactId>
  <version>${selenium.java}</version>
  <scope>test</scope>
</dependency>
```

```
<dependency>
  <groupId>org.testng</groupId>
  <artifactId>testng</artifactId>
  <version>${testng}</version>
  <scope>test</scope>
</dependency>
```

Для підтримки патерну Page Object ми використаємо Page Factory, який являє собою розширення Page Object. Ми будемо використовувати Page Factory для ініціалізації веб-елементів, які визначені в класах веб-сторінки:

```
public BasePage(WebDriver driver) {
    this.driver = driver;
    PageFactory.initElements(driver, this);
}
```

Page Factory ініціалізує кожний веб-елемент сторінки із використанням анотації @FindBy. За допомогою цієї анотації ми визначаємо стратегію пошуку елемента, а також необхідну інформацію для його ідентифікації:

```
@FindBy(xpath = "//button[@name='Login1']")
private WebElement logInButton;
```

Також із використанням анотацій @BeforeMethod та @AfterMethod прописуємо необхідні налаштування, які будуть відповідати відповідно початку та завершенню кожного тестового сценарію.

IV. РЕЗУЛЬТАТИ

Основними складовими створених класів буде веб-драйвер (завдяки якому ми взаємодіємо з браузером), методи (які ідентифікують дії користувача на певній сторінці), веб-елементи сторінки з відповідними локаторами з модифікатором доступу private та анотацією @FindBy.

V. ОБГОВОРЕННЯ

Для того, щоб написати свій тестовий фреймворк на базі Page Object та Page Factory необхідно виконати аналіз веб-додатку, який ми тестуємо. Потрібно проаналізувати тестові дані, які ми будемо використовувати. Для випадку, де їхня кількість буде великою чи є ймовірність, що вона буде збільшуватися, буде доречно додати їх у property-файл (папка resources).

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У подальшому з метою безперервної інтеграції наших тестів, як приклад у Jenkins, можна у property-файли додати середовища тестування нашого веб-додатку, що дозволить автоматизувати частину процесу розробки програмного забезпечення.

VII. ВИСНОВКИ

Page Object та Page Factory спрощують моделювання веб-сторінок в Selenium. Класи Page Object можна використовувати повторно для всього набору тестів, що дає можливість інтегрувати розроблені тести для своїх проєктів на ранніх стадіях, зберігаючи легкі та прості процедури процесу тестування. Це дозволяє адаптувати свій набір тестових сценаріїв до змін на проєкті без затрати особливих зусиль та часу.

ЛІТЕРАТУРА

- [1] Erich Gamma, Richard Helm, Ralph Johnson, John Vissdes, "Design Patterns: Elements of Reusable Object-Oriented Software". United States: Addison-Wesley, 1994. 395 p.
- [2] Seretta Gamba, Dorothy Graham, "A Journey through Test Automation Patterns: One team's adventures with the Test Automation Patterns". United States: CreateSpace Independent Publishing Platform, 2018. 364 p.
- [3] Anton Angelov, "Design Patterns for High-Quality Automated Tests: High-Quality Test Attributes and Best Practices". United States: Kindle Edition, 2021. 348 p.

Постановка задачі розробки програмного забезпечення для оцінювання ефективності протиепідемічних заходів та прогнозування зміни відсотка інфікованих та перенесених хвороб у тяжкій формі

Олександр Мельников, Дмитро Козуб

*Кафедра інтелектуальних систем прийняття рішень
Донбаська державна машинобудівна академія
Краматорськ, Україна*

Анотація—Розглянуто проблему вивчення впливу протиепідемічних заходів на відсоток інфікованих пандемічним захворюванням. Описано створений раніше додаток, який дозволяє здійснювати оброблення інформації щодо вакцинованих осіб. Сформульовано задачу розробки спеціалізованого програмного забезпечення – системи підтримки прийняття рішень для оцінювання ефективності протиепідемічних заходів та прогнозування зміни відсотка інфікованих та перенесених хвороб у тяжкій формі. Наведено приклад розв’язання задачі у середовищі аналітичного пакета.

Ключові слова—COVID-19, Пандемія, Протиепідемічні заходи, Програмне забезпечення, Система підтримки прийняття рішень.

I. ВСТУП

COVID-19 (SARS-CoV-2) є небезпечним захворюванням, яке може протікати як у формі гострої респіраторної вірусної інфекції легкого перебігу, так і у важкій формі. Вакцинація є одним з найкращих методів для захисту від COVID-19. Хоч він не дає 100% гарантії захисту від COVID-19, але забезпечує проходження хвороби у легкій формі, без ускладнень та гарантує швидке одужання [1-3]. Створення та використання спеціалізованого програмного забезпечення для оцінювання ефективності протиепідемічних заходів та прогнозування зміни відсотка інфікованих та перенесених хвороб у тяжкій формі може здійснити суттєвий вплив щодо покращення ситуації у цьому напрямку.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Наразі немає спеціалізованих програмних засобів або онлайн сервісів, які забезпечували б розв’язання поставленої задачі. Автори спроектували [4] та реалізували [5] додаток – програмне забезпечення для моніторингу вакцинованих студентів у навчальному закладі, але він не розв’язував задачі вивчення ефективності протиепідемічних заходів та прогнозування кількості хворих або одужалих.

III. МЕТОДИ

Основна ідея: є статистичні дані (щомісячні) про збільшення відсотка інфікованих під час епідемії (пандемії) та тих, хто переносить хворобу у тяжкій формі, а також переліку протиепідемічних заходів, що застосовуються в цей період у даному регіоні [6-7]. (Під поняттям «регіон» мається на увазі місто, район або область, керівництво якого наділене повноваженнями вводити або скасовувати будь-які протиепідемічні заходи, і переміщення за

межі якого можна суттєво обмежити). Завдання – оцінити ефективність вжитих протиепідемічних заходів і здійснити прогнозування зміни відсотка інфікованих і тих, хто переносить хворобу у тяжкій формі.

Можна виділити такі вхідні фактори нашої прогнозовної моделі:

- обов'язковий «масковий режим» (Masks);
- введення карантину, тобто скасування масових заходів, встановлення антисептиків у всіх адміністративних закладах тощо (Quarantine);
- введення дистанційного навчання у навчальних закладах (Distance_Learning);
- можливість вільного вакцинування (Vaccine_optional);
- запровадження обов'язкового вакцинування (Vaccination_is_mandatory);
- відсоток вакцинованих (Percentage_of_Vaccinated);

Вихідні фактори:

- зміна відсотка інфікованих (Infected);
- зміна відсотка тих, хто переносить хворобу у тяжкій формі (Severe_cases).

Частина факторів є біноміальними, а частина – числовими безперервними. Також слід зазначити, що для другої задачі (тобто розрахунку змінення відсотка тих, хто переносить хворобу у тяжкій формі) зміна відсотка інфікованих може стати додатковим вхідним фактором.

Приклад вхідних даних наведено на рис. 1.

Region	Masks	Quarantine	Distance	Vaccin	Vaccination_is_m	Percentag	Infected	Severe_cases
Kyivskyi	0	1	0	0	0	11,98	42,12	17
	1	1	1	0	0	20,15	38,62	13,8
	1	1	1	1	0	44,22	24,13	9,3
	1	1	1	1	1	47,28	13,53	5
Donetsk	0	1	0	0	0	7,2	29,74	16,1
	1	1	1	0	0	9,06	18,12	12,1
	1	1	1	1	0	13,67	11,62	8,7
	1	1	1	1	1	15,59	5,49	5,9
Dnipro	0	1	0	0	0	18,74	36,12	15,9
	1	1	1	0	0	36,12	24,35	14,1
	1	1	1	1	0	40,48	15,24	8,1
	1	1	1	1	1	44,49	10,32	6,2
Kharkiv	0	1	0	0	0	13,86	37,14	15,9

Рисунок 1 – Приклад вхідних даних

IV. РЕЗУЛЬТАТИ

Далі задачу прогнозування можна розв'язувати різними методами. Для прикладу наведемо рішення в середовищі Deductor Studio Lite методом штучних нейронних мереж. Архітектуру мережі (багатошаровий перцептрон 5x5x2, 5 нейронів в одному прихованому шарі) зображено рис. 2., результати розрахунків – на рис. 3. Використання моделі для прогнозування (рис. 4) показує, що введення обов'язкового вакцинування знижує зростання відсотка інфікованих на 1,5%, а відсотка хворих на тяжку форму – майже вдвічі.

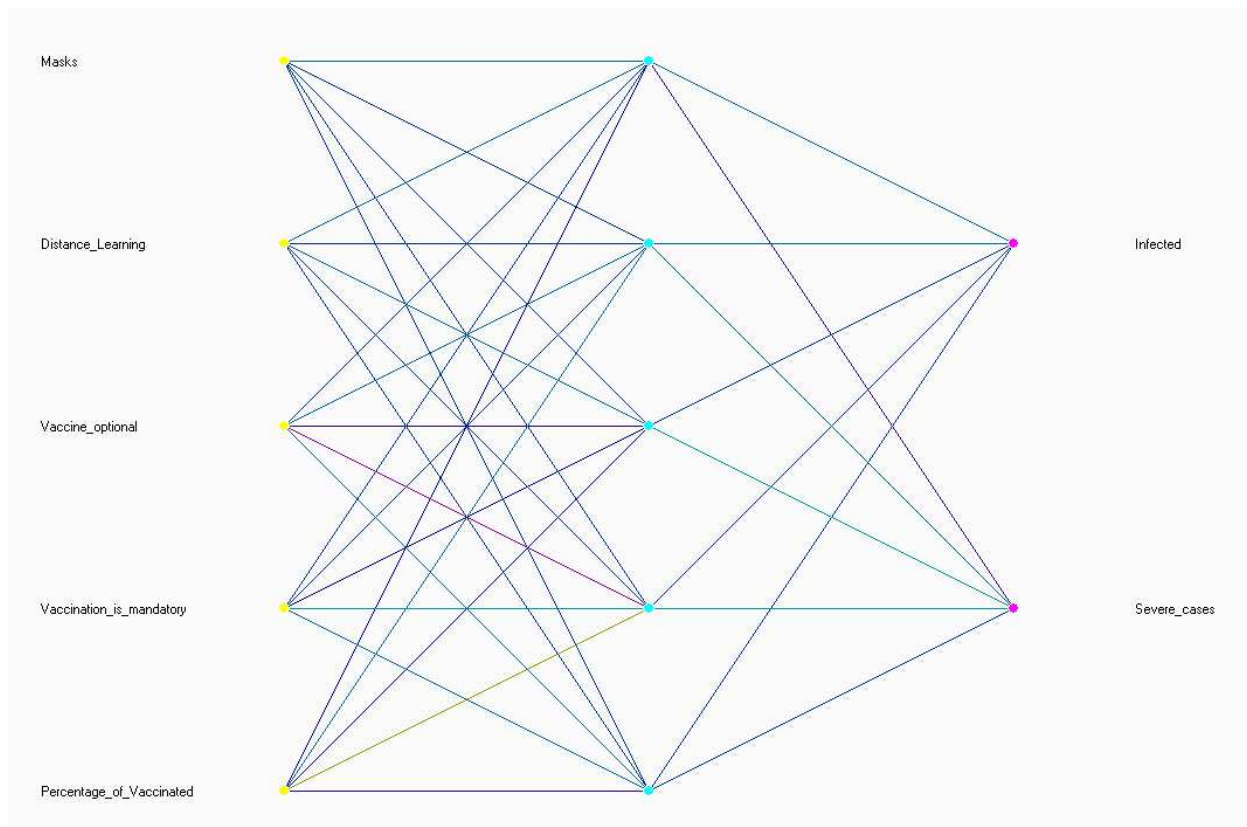


Рисунок 2 – MLP 5x5x2

Region	Masks	Quarantine	Distance_Learning	Vaccine_optional	Vaccination_is_mandatory	Percentage_of_Vaccinated	Infected	Infected_OUT	Severe_cases	Severe_cases_OUT
Kyivskiy	0	1	0	0	0	11,98	42,12	40,0011845051581	17	16,2594789214857
	1	1	1	0	0	20,15	38,62	31,5073371423768	13,8	12,7079860174533
	1	1	1	1	0	44,22	24,13	22,2751043213315	9,3	9,07407487381311
	1	1	1	1	1	47,28	13,53	13,7672176085666	5	5,83285699195621
Donetsk	0	1	0	0	0	7,2	29,74	29,7524126142231	16,1	16,1119646870169
	1	1	1	0	0	9,06	18,12	18,0910727227571	12,1	12,160797671705
	1	1	1	1	0	13,67	11,62	11,6165701472529	8,7	8,70052889732756
	1	1	1	1	1	15,59	5,49	5,6950006105068	5,9	5,86440295309101
Dnipro	0	1	0	0	0	18,74	36,12	39,4742872189235	15,9	16,6048910945423
	1	1	1	0	0	36,12	24,35	22,4997959748433	14,1	13,1905857015608
	1	1	1	1	0	40,48	15,24	21,7699517800451	8,1	9,07756215042054
	1	1	1	1	0

Рисунок 3 – Результати розрахунків

Рисунок 4 – Використання моделі для прогнозування

V. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Далі треба здійснити програмну реалізацію усіх сформульованих завдань.

VI. ВИСНОВКИ

Було сформульовано задачу створення спеціалізованого програмного забезпечення для оцінювання ефективності протиепідемічних заходів та прогнозування зміни відсотка інфікованих та перенесених хвороб у тяжкій формі, яке може здійснити суттєвий вплив щодо покращення ситуації у цьому напрямку.

ЛІТЕРАТУРА

- [1] COVID-19 [Електронний ресурс]. Режим доступу: <https://en.wikipedia.org/wiki/COVID-19>. Дата звернення: Жов. 28, 2022.
- [2] COVID-19 pandemic in Ukraine [Електронний ресурс]. Режим доступу: https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Ukraine. Дата звернення: Жов. 28, 2022.
- [3] Вакцинація проти COVID-19 в Україні [Електронний ресурс]. Режим доступу: https://uk.wikipedia.org/wiki/Вакцинація_проти_COVID-19_в_Україні. Дата звернення: Жов. 28, 2022.
- [4] Д. С. Козуб, О. Ю. Мельников, «Постановка задачі розробки програмного забезпечення для відстежування вакцинованих студентів в навчальному закладі», *Молодь і наука: виклики та перспективи: збірник тез наукової конференції молодих вчених 16 грудня 2021р.*, Краматорськ: Донецька обласна державна адміністрація, Рада молодих вчених при Донецькій облдержадміністрації, 2021, с. 144-145
- [5] О. Ю. Мельников, Д. С. Козуб, «Розробка програмного забезпечення для моніторингу вакцинованих студентів у навчальному закладі», *Автоматизація та комп'ютерно-інтегровані технології у виробництві та освіті: стан, досягнення, перспективи розвитку: матеріали Всеукраїнської науково-практичної Internet-конференції*, Черкаси, 2022, с. 57-59
- [6] Коронавірус в Україні - Статистика [31.10.2022] - Карта заражень, графіки [Електронний ресурс]. Режим доступу: <https://index.minfin.com.ua/ua/reference/coronavirus/ukraine>. Дата звернення: Жов. 31, 2022.
- [7] Вакцинація від коронавірусу в Україні - Статистика 23.02.2022 - Мапа, графіки [Електронний ресурс]. Режим доступу: <https://index.minfin.com.ua/ua/reference/coronavirus/vaccination/ukraine/>. Дата звернення: Жов. 31, 2022.

Нелінійні нормальні моди маятникової системи під дією магнітних збуджень

Юлія Сурганова, Юрій Міхлін

Кафедра прикладної математики

Національний технічний університет «Харківський політехнічний інститут»

Харків, Україна

Анотація—Розглянуто динаміку двох пов'язаних маятників при магнітному збудженні. Інерційні компоненти маятників істотно розрізняються, і за малий параметр обирається співвідношення мас маятників. Наближення Паде використовуються для опису магнітного збудження з хорошою точністю. Метод малих параметрів використовується для побудови двох нелінійних нормальних мод коливань (або NNMs, як прийнято в відповідній літературі), одна з яких є зв'язаною, а друга - локалізованою. Вивчено вплив співвідношення мас та інших параметрів системи на NNMs як для малих, так і не малих початкових кутів маятників.

Ключові слова: Пов'язані маятники, Магнітне збудження, Нелінійні нормальні моди коливань.

I. ВСТУП

Математичні або фізичні маятники є важливими моделями, що представляють типову нелінійну динаміку різноманітних нелінійних систем. Вони можуть проявляти як регулярну, так і істотно складну поведінку при різних видах зовнішніх впливів. Зокрема, у кількох останніх публікаціях представлені теоретичні та експериментальні дослідження динаміки двох маятників при електромагнітному збудженні [1–4].

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

В теорії нелінійних коливань, а також у прикладних задачах проблема локалізації коливань є дуже важливою і досліджувалася в останні десятиліття в численних публікаціях. Наявність зовнішніх впливів, зокрема, магнітних сил, призводить до значного ускладнення аналізу коливань нелінійних систем. Метою даної роботи є аналітичне і чисельне дослідження NNMs в системі, що складається з двох пов'язаних маятників, під впливом електромагнітних сил. Моделі двох пов'язаних маятників при електромагнітному збудженні і взаємодії кожного маятника з магнітом розглянуті в публікаціях [1-3] для випадку, коли маси маятників рівні. У даній роботі розглянуто випадок, коли маси істотно розрізняються. В цьому випадку можлива локалізація енергії вібрації. Використовуються методи теорії нелінійних нормальних форм коливань, що представлені в численних публікаціях відносно нелінійних систем з кількома степенями свободи (див., зокрема, огляди [5,6]).

III. МЕТОДИ

Основні рівняння динаміки системи представлено як

$$\begin{cases} \varepsilon I \ddot{\varphi}_1 = \varepsilon M_{mag_1} + \varepsilon M_{D1}(\varphi_1, \varphi_2) + \varepsilon M^{(g)}(\varphi_1) + M^{(k)}(\varphi_1, \varphi_2), \\ I \ddot{\varphi}_2 = \varepsilon M_{mag_2} + \varepsilon M_{D2}(\varphi_1, \varphi_2) + M^{(g)}(\varphi_1) + M^{(k)}(\varphi_1, \varphi_2). \end{cases} \quad (1)$$

Тут I - момент інерції, $M_{mag_{1,2}}$ представляють електромагнітні сили, $M_{D_{1,2}}$ - це моменти сил тертя, $M_{D1} = -C_1 - C_e(\varphi_1 - \varphi_2)$ - це моменти сил демпфірування, $C_1\dot{\varphi}_1$ та $C_2\dot{\varphi}_2$ - моменти опору руху в середовищі, $C_e(\dot{\varphi}_1 - \dot{\varphi}_2)$ і $C_e(\dot{\varphi}_2 - \dot{\varphi}_1)$ - моменти сил демпфірування, що стверені пружним елементом, $M^{(g)}, M^{(k)}$ - моменти гравітаційних та пружних сил,

відповідно, $M^{(g)} = -mgs \sin \varphi_{1,2}$, $M^{(k)} = -k_l(\varphi_1 - \varphi_2)$, $r = mgs$, причому синуси розкладаються в скорочені ряди Тейлора, а саме, $\sin(\varphi) \cong \varphi - \frac{\varepsilon}{6} \cdot \varphi^3$, s – відстань від центру мас до осі обертання, m - маса більшого маятника, ε - малий параметр, що характеризує відношення мас маятників, k_l - коефіцієнт, що характеризує жорсткість згину пружного елемента. Для опису магнітної взаємодії було використано апроксимацію Паде. Параметри обраної моделі було знайдено за допомогою процедури методу найменших квадратів щоб найкращим чином задовольнити експериментальним даним, представленим в [1]. Порівняння отриманої апроксимації та експерименту представлено на Рис. 1.

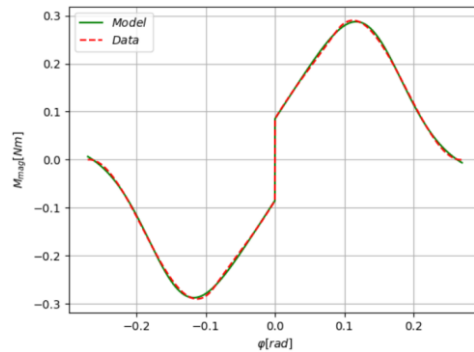


Рисунок 1. Порівняння апроксимації магнітного моменту з експериментальними даними.

Одночастотна мода коливань визначається методом малого параметру. Аналітичні результати, отримані у двох наближеннях за малим параметром, були зіставлені з чисельним моделюванням вихідної системи, яке було отримано з використанням оригінальних програм, написаних на мові програмування Python. Зокрема, чисельне рішення в часовій області визначається методом Рунге–Кутти 4-го порядку. Порівняння проводиться при початкових умовах, які задаються аналітичним рішенням.

IV. РЕЗУЛЬТАТИ

Хоча основна увага в цій роботі була приділена проблемі впливу параметрів системи на нелінійні нормальні режими вібрації, проте також визначено основні типи поведінки системи поблизу режимів вібрації. Зроблено спробу описати основні типи поведінки системи, близькі до вказаних нормальних мод коливань. Зокрема, виділено резонансні режими руху системи. Представлено результати дослідження пов'язаних і локалізованих режимів вібрації як при відсутності магнітного впливу, так і при наявності його. Як приклад, на Рис. 2 показана поведінка локалізованої форми коливань при малих початкових кутах маятників (Рис. 1а), та при не малих кутах (Рис. 2б).

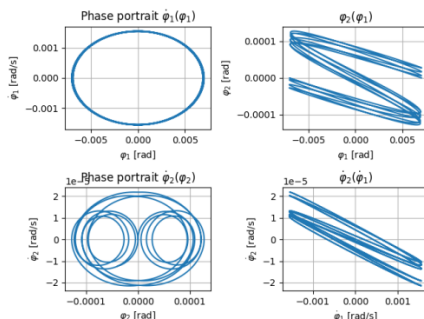


Рис. 2а

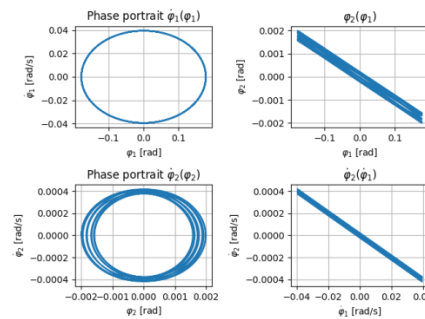


Рис. 2б

Рисунок 2. Траєкторії локалізованої форми коливань в фазовому та конфігураційному просторах при малих початкових кутах маятників (Рисунок 2а) та при не малих кутах (Рисунок 2б)

V. ОБГОВОРЕННЯ

Оскільки при чисельному моделюванні немає можливості використовувати початкові значення режимів вібрації з абсолютною точністю, то у результаті неможливо отримати

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

точне одночастотне періодичне рішення. Однак в околиці NNMs можна отримати як рішення, дуже близькі до режиму одночастотних коливань, коли вплив другої частоти невеликий, так і рішення, далекі від одночастотних. Авторами було використано термін «блукання» (або) для опису траєкторій квазіперіодичних рухів в фазовому та конфігураційному просторах, які можуть бути як близькими, так і віддаленими щодо відповідних коливальних режимів [5].

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Чисельне моделювання, включаючи розрахунок частотного спектру та аналіз стійкості розглянутих ННР при зміні параметрів системи і початкових кутів, дозволять виявити як регулярну поведінку у формі ННР з, можливо, деякими невеликими відхиленнями поблизу траєкторії ННР, так і нерегулярну поведінку.

VII. ВИСНОВКИ

Отримано, що режим пов'язаних коливань, як із зменшеними, так із збільшеними магнітними впливами, існує на всьому діапазоні початкових відхилень кутів маятників, але при кутах менше за $4-5^\circ$, помітні блукання траєкторії поблизу моди коливань, а синфазні і відповідні квазіперіодичні рішення з'являються поряд з цим режимом. Руйнування режиму нормальної моди коливань відбувається при великому значенні коефіцієнта співвідношення мас маятників. Інші параметри системи, наприклад, коефіцієнт зв'язку та відстань від центру мас маятника до осі обертання, не приводять до руйнування цього режиму (за винятком резонансної ситуації). Локалізована форма існує при початковому значенні кута меншого маятника більше за $10-16^\circ$. При початкових умовах, починаючи приблизно з 10° , при невеликому магнітному ефекті збільшення малого параметра (тобто збільшення відношення мас) призводить до стійкого локалізованого режиму, коли всі блукаючі траєкторії зникають; у той же час при значному магнітному впливі збільшення малого параметра спочатку призводить до стабільного локалізованого режиму, що усуває всі блукання траєкторії, а потім ці блукання виникають знову, і локалізація перестає спостерігатися. При великих кутах із збільшенням зчеплення між маятниками з'являється стабільна локалізована форма. А при малих початкових кутах маятників збільшення коефіцієнта зв'язку для всіх розглянутих типів магнітного впливу не усуває блукання, тому локалізована мода себе не проявляє. При дуже малому початковому куті маятників (незалежно від величини зчеплення), зі збільшенням відстані від центру мас маятника до осі обертання блукаючі траєкторії квазіперіодичних режимів не зменшуються. У той час як при великих кутах (навіть якщо з'єднання невелике) збільшення такої відстані майже повністю усуває усі блукання, і локалізована форма стає більш визначеною.

ЛІТЕРАТУРА

- [1] A. Wijata, K. Polczyński, J. Awrejcewicz, "Theoretical and numerical analysis of regular one-side oscillations in a single pendulum system driven by a magnetic field", *Mech. Syst. Signal Process.* 150 (2021) 107229, doi: 10.1016/j.ymsp.2020.107229
- [2] K. Polczyński, S. Skurativskiy, M. Bednarek, J. Awrejcewicz, "Nonlinear oscillations of coupled pendulums subjected to an external magnetic stimulus, *Mech. Syst. Signal Process.*". 154 (2021) 107560, doi: 10.1016/j.ymsp.2020.107560
- [3] S. Skurativskiy, K. Polczyński, M. Wojna, J. Awrejcewicz, "Quantifying periodic, multi-periodic, hidden and unstable regimes of a magnetic pendulum via semi-analytical, numerical and experimental methods", *J. Sound Vib.* 524 (2022) 116710, doi: 10.1016/j.jsv.2021.116710
- [4] Yu.E. Surganova, Yu.V. Mikhlin, "Localized and non-localized nonlinear normal modes in a system of two coupled pendulums under a magnetic field", *Int. J. of Non-Linear Mech.* 147 (2022) 104182, doi: /10.1016/j.ijnonlinmec.2022.104182
- [5] Yu. V. Mikhlin, K. V. Avramov, "Nonlinear normal modes for vibrating mechanical systems. Review of theoretical development", *Appl. Mech. Rev.* 63 (6) (2010) 060802, doi: 10.1115/1.4003825
- [6] C. E.N.Mazzilli, P. B.Gonçalves, G. R.Franzini. "Reduced-order modelling based on non-linear modes". *Int. J. of Mech. Sciences.*214, 2022, 106915, doi: /10.1016/j.ijmecsci.2021.106915

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

Дослідження використання паралельних технологій у високо-навантажувальних додатках

Ярослав Марчук

*Кафедра прикладної математики
Львівський національний університет імені Івана Франка
Львів, Україна*

Сучасні корпоративні системи включають в себе бекенд додатки, які використовують REST API. Досить часто додатки REST API містять в собі розрахунки складних алгоритмів, які потребують величезної кількості ресурсів. Для оптимізації виконання застосовують різні технології паралельних розрахунків. У даній роботі виконано порівняння швидкодії REST API додатків реалізованими паралельними технологіями MPI (Message Passing Interface – інтерфейс передачі повідомлень), goroutine, та підходом реактивного програмування. Реалізація на Python була використана як найбільш найпопулярна та найбільш поширена мова програмування. Go використана як спеціалізована мова програмування для бекенд додатків. Підхід goroutine був використаний для Go. Реалізація на Java проведена разом з підходом реактивного програмування, що включає в себе асинхронність та потокову обробку даних. Elixir – функціональна, розподілена мова програмування, яка працює у віртуальній машині Erlang. Особливістю Elixir є можливість програмування на рівні окремих процесів, комунікація між якими відбувається з допомогою MPI.

Ключові слова—REST API, Python, Golang, goroutines, Java, Elixir, MPI.

I. ВСТУП

У сучасному світі ми знайдемо багато проблем, які вимагають значного та ефективного використання обчислювальних ресурсів. Більшість додатків, які потребують продуктивності, традиційно належать до наукової сфери. Але, згідно прогнозів, програми штучного інтелекту і машинного навчання стануть переважними користувачами великомасштабних обчислень.

Звичайні програми залишають невикористаними обчислювальні можливості сучасних комп'ютерів. Паралельні обчислення є ключем до розкриття потенціалу ресурсів комп'ютера. Однією з нових сфер впливу паралельних обчислень є енергоефективність.

Паралельні обчислення — це виконання багатьох операцій в один момент часу. Повне використання паралельних обчислень не відбувається автоматично. Це вимагає певних зусиль від розробника програмного забезпечення. По-перше, ми повинні визначити й виявити потенціал паралелізму в програмі. Потенційний паралелізм або паралелізм означає, що ми підтверджуємо, що безпечно виконувати операції у будь-якому порядку, коли системні ресурси стають доступними. І для паралельних обчислень існує додаткова вимога: ці операції мають виконуватися одночасно. Щоб це сталося, ми також повинні належним чином використовувати ресурси для одночасного виконання цих операцій.

Майбутнє паралельне. Збільшення продуктивності серійного процесора зупинилося, оскільки конструкції процесорів досягли меж мініатюризації, тактової частоти, потужності та навіть тепла.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Паралельні обчислення створюють нові проблеми, яких немає в послідовному світі. Паралельні обчислення потребують поєднання розуміння апаратного забезпечення, програмного забезпечення та паралельності для розробки програми. У послідовних обчисленнях усі операції прискорюються зі збільшенням тактової частоти. На відміну, в паралельних обчисленнях, нам потрібно трохи подумати та змінити наші програми, щоб повністю використовувати паралельне апаратне забезпечення. Нам потрібно змінити процеси мислення, щоб адаптуватися до додаткових складнощів паралельного виконання, але з часом це дає нам хороші результати.

Розроблені додатки базуються на реалізації паралельних особливостей кожної використаної мови програмування.

III. МЕТОДИ

Python був обраний через його широке використання та простоту вивчення новачками. Ключову роль популярності мови демонструє величезна кількість розроблених бібліотек, а також постійне і неухильне вдосконалення мови та бібліотек.

REST API реалізований мовою Python у фреймворку Django. Головною перевагою використання Django для цієї роботи є можливість керування запитамі RESTful та широка документація. Цей фреймворк налічує понад 1 мільйон користувачів. Це забезпечує постійні оновлення фреймворку для підтримки популярності та потреби в таких інструментах.

Потоки мають мінімальний розмір пам'яті; зазвичай це розмір МБ (2 МБ для Linux). Мінімальний розмір встановлює певні обмеження на створення програми для нового потоку — якщо кожен потік має принаймні кілька МБ, 1000 потоків займатимуть принаймні кілька ГБ пам'яті. Го вирішує ці проблеми за допомогою конструкції, подібної до потоків, але це обробляється середовищем виконання мови, а не ОС. Розмір goroutine у пам'яті становить три порядки величини (2 Кб на goroutine), що означає, що мінімальне використання пам'яті 1000 goroutine можна порівняти з використанням пам'яті одним потоком.

Реактивне програмування означає програмування реакцій на події, які надсилаються асинхронно. Це означає програмування компонента, який отримує та створює повідомлення. Між ними можна додавати перетворення, фільтри або компоненти, які створюють або отримують повідомлення.

Бібліотека RxJava використовувалася під час написання реалізації REST API на Java. RxJava - це реалізація віртуальної машини Java бібліотеки ReactiveX для створення асинхронних і керованих подіями додатків з використанням послідовностей, що спостерігаються.

Будівельними блоками RxJava є Observables та Subscribers. Об'єкти, що спостерігаються, використовуються для публікації елементів, а передплатники використовуються для споживання цих елементів.

Важливість стандарту інтерфейсу передачі повідомлень (MPI) полягає в тому, що він дозволяє програмі отримувати доступ до додаткових обчислювальних вузлів і, таким чином, запускати все більші й більші проблеми, додаючи більше вузлів до симуляції. Ім'я передачі повідомлень означає можливість легко надсилати повідомлення від одного процесу до іншого. MPI є всюдисущим у сфері високопродуктивних обчислень. У багатьох галузях науки використання суперкомп'ютерів передбачає впровадження MPI.

IV. РЕЗУЛЬТАТИ

Для зручного навантажувального тестування серверів, що розробляються, використовується популярний інструмент Bombardier.

У таблиці 1 показані результати виконання 3000 запитів по 5 з'єднанням з максимальною затримкою 10 секунд кожного реалізованого сервера.

bombardier -c 5 -n 3000 -t 10s server_URL

	Python	Golang	Java	Elixir
Загальний час, хв	25,01	12,2	14,1	0,04
Запитів/сек	2,79	5,32	4,97	752
Затримка, сек	3,71	1,82	2,03	0,00687
Пропускна здатність, КБ/сек	302	917,2	826,82	2205

Таблиця 1

Розглянемо графік загального часу виконання запиту сервером (рис.1), затримки відповіді від сервера (рис.2), кількості запитів за секунду (рис.3) та пропускної спроможності сервера (рис.4).



Рис.1 Гістограма часу виконання запитів

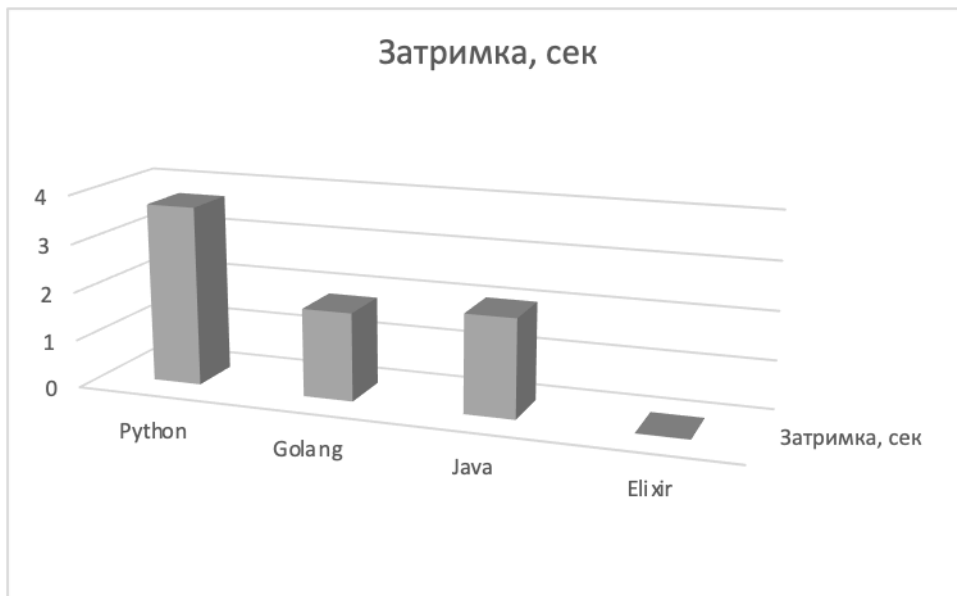


Рис.2 Гістограма затримки надсилання відповідей

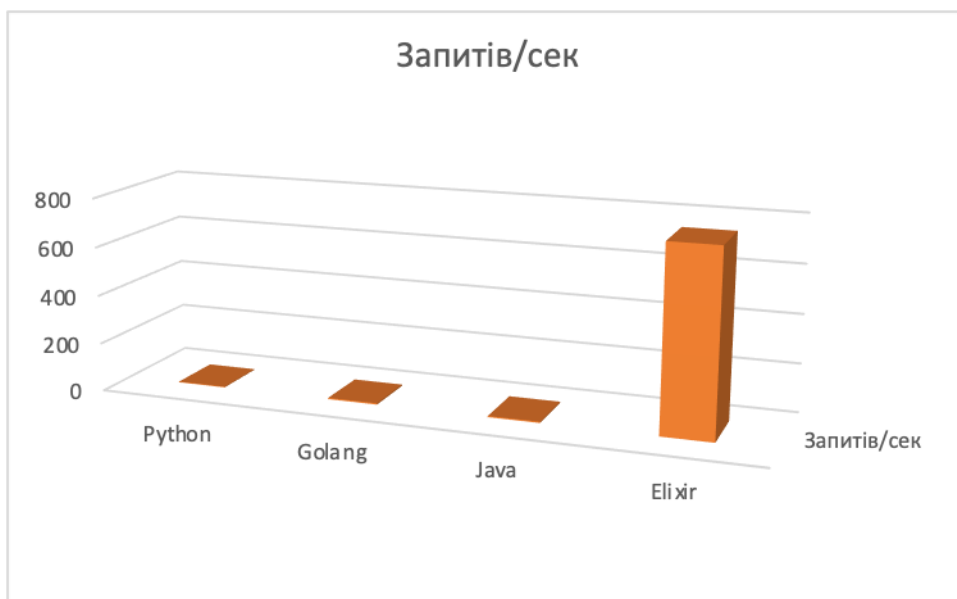


Рис.3 Гістограма кількості запитів за секунду



Рис.4 Гістограма пропускної здатності серверів

У таблиці 2 показані результати виконання 7000 запитів з 10 з'єднанням з максимальною затримкою 10 секунд для кожної реалізації сервера:

`bombardier -c 10 -n 7000 -t 10s server_URL`

	Python	Golang	Java	Elixir
Загальний час, хв	24,93	16,7	17,64	0,19
Запитів/сек	2,68	2,69	2,34	286,3
Затримка, сек	3,71	2,76	3,42	0,03531
Пропускна здатність, КБ/сек	288	519,2	361,9	856,19

Таблиця 2



Рис.1 Гістограма часу виконання запитів

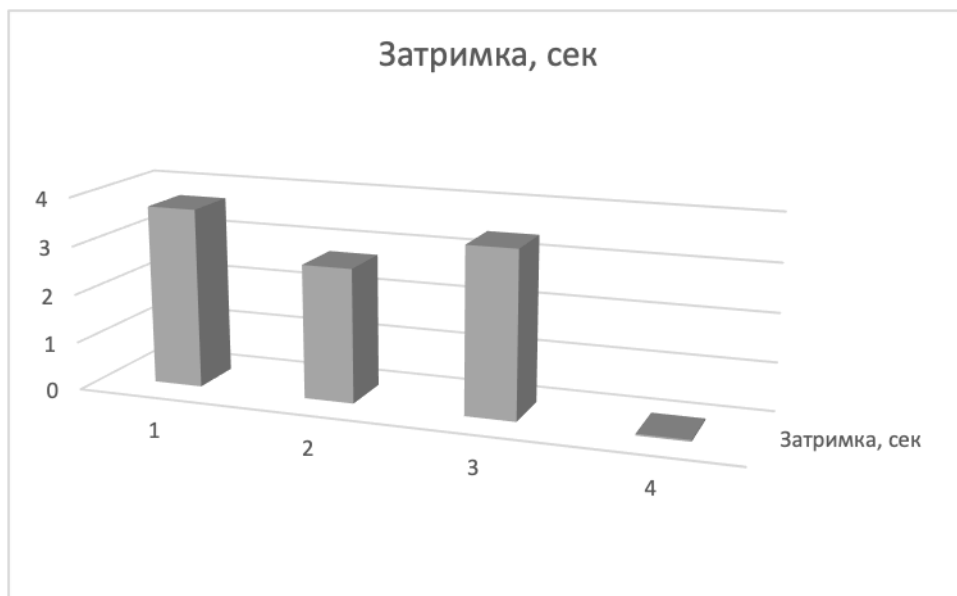


Рис.2 Гістограма затримки надсилання відповідей

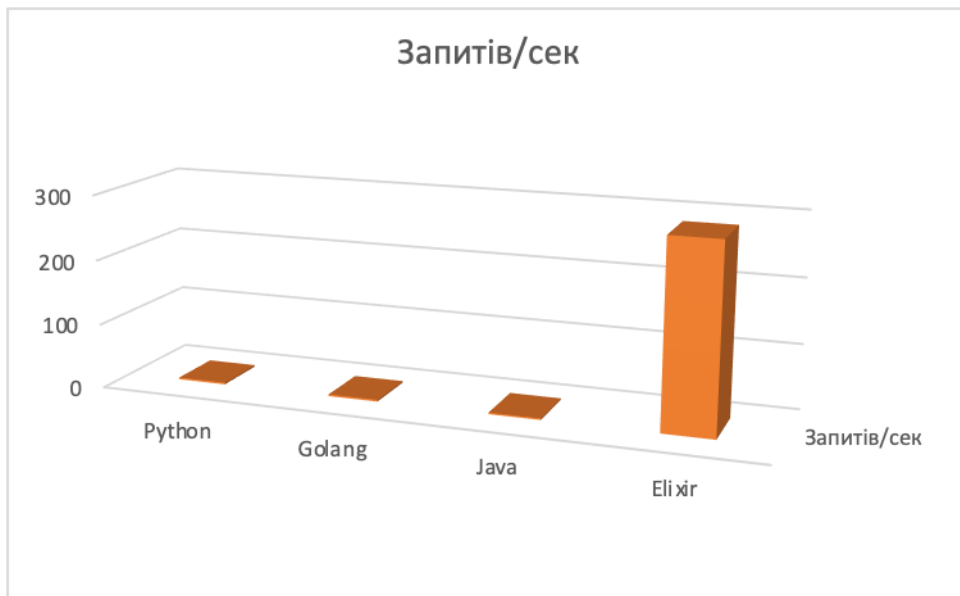


Рис.3 Гістограма кількості запитів за секунду



Рис.4 Гістограма пропускної здатності серверів

V. ОБГОВОРЕННЯ

В даній роботі розрахунки проводились не на спеціалізованих надзвичайно швидких апаратних системах та не на оптимізованих операційних системах під високонавантажувальні додатки. Ці обмеження були зумовлені відсутності доступу до таких систем.

Всі розрахунки проводились на одному і тому ж сервері. Навіть в цих умовах можна побачити, що отримані результати показують, що майбутнє високонавантажувальних додатків за вузько спеціалізованими інструментами. Система на Erlang з вбудованим MPI показала найвищу швидкодію. Це дійсно підтверджується прикладами використання Erlang для бекенд додатків в Приватбанку та в електронній системі охорони здоров'я України(eHealth).

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Для подальшого розвитку даного дослідження буде використано хмарні технології з балансуванням навантаження. З використанням цього підходу буде змога масштабувати додатки. Балансувальник навантаження підтримує як вхідні, так і вихідні напрямки потоків даних. Балансувальник навантаження забезпечує низьку затримку та високу пропускну здатність, а також масштабує до мільйонів потоків для наступних протоколів TCP і UDP.

VII. ВИСНОВКИ

На підставі вимірювання швидкості розроблених REST-додатків можна зробити висновок, що універсальна мова програмування забезпечує практично однакову швидкість з невеликим відхиленням. Golang наголошує на мережевому навантаженні, тому в порівнянні з реалізацією на Python, Golang з горутинами показав кращу продуктивність. Java у поєднанні з підходом реактивного програмування не відповідав менш як на 1% запитів. Це пов'язано з тим, що надбудови реактивного підходу потребують більше ресурсів. Функціональний підхід до реалізації REST API засобами Elixir із паралельною обробкою даних з допомогою MPI показав найкращі результати.

ЛІТЕРАТУРА

- [1] M. Masse, "REST API Design Rulebook", Design Rulebook: O'Reilly Media, 2011, pp. 23–78
- [2] C. Gaston, "Django RESTful Web Services", Packt Publishing, 2018, pp. 178-202
- [3] A. Ravindran, "Django Design Patterns and Best Practices", *Second Edition*, Packt Publishing, 2018, pp. 126-179
- [4] A. Saha, "Practical Go", Wiley, 2021, pp. 281-284
- [5] M. Tsoukalos, "*Mastering Go*", Second Edition, Packt Publishing, 2021, pp. 208-216
- [6] T. Nurkiewicz, B. Christensen, "*Reactive Programming with RxJava*", O'Reilly Media, 2016, pp.438-451
- [7] S. Jurich, "Elixir in Action", Second edition, Manning Publications, 2019, pp.34-68
- [8] S. Gospodinov, "Concurrent Data Processing in Elixir," Pragmatic Bookshelf, 2021, pp.96-104

Garbage classification for mobile devices based on MobileNet

Ivan Rzhpetskyi, Anastasiya Troynina

*Institute of computer science
Odesa Polytechnic National University
Odesa, Ukraine*

Abstract—The paper proposes research, the purpose of which is to build a neural network for garbage classification, which would be able to work on embedded devices that differ in a smaller amount of RAM and computing capabilities. As a result of the work, a neural network was created that can classify 12 types of household garbage and works on mobile devices.

Keywords—*Software Engineering, Computer Science, Deep Learning.*

I. INTRODUCTION

According to OECD report, only 9% of all plastic waste is currently recycled. Also, based on the same report, people produce twice as much plastic as was produced 20 years ago [1]. It's data about plastic, but there are other types of waste, like glass, metal, plastic, paper. To keep balance, we need to increase percentage of recycled garbage every year. The problem is that sorting machines are expensive and that's makes the problem harder. The developed algorithm is intended for decreasing costs of sorting by resolving classification task and detecting 12 classes of waste right in embedded systems. For example, this system can be connected to camera, installed in a garbage bin, and signalize that people are sorting garbage incorrectly.

II. RELATED WORK

Current work was inspired by research made by Kang, Zhuang & Yang, Jie & Li, Guilan & Zhang, Zeyi "An Automatic Garbage Classification System Based on Deep Learning" [2]. Based on this article, the waste classification problem was solved using ResNet architecture and could classify 14 types of garbage with accuracy equals to 0.999. The problem is that these networks have a big number of parameters, which makes hard to work on embedded systems with low amount of RAM and weak CPU. That was the main reason for future research.

III. METHODOLOGY

One of the biggest bottlenecks in machine learning is not very big number of datasets, that's why for most cases dataset should be gathered from different sources, resulting in different image quality. Training data for this research was taken from Kaggle website [3]. This dataset consists of images from web, "Garbage Classification" dataset made by CCHANG [4] and "Clothing dataset" [5]. Manually discovering images from dataset, we can notice that part of the images was programmatically improved, part remain the same as they were. By programmatically improved images I mean images with adjusted color balance, removed background. This can be seen for battery category. But most of the images are real, especially in shoes and clothes class. Images distribution and available classes can be seen on fig.1.

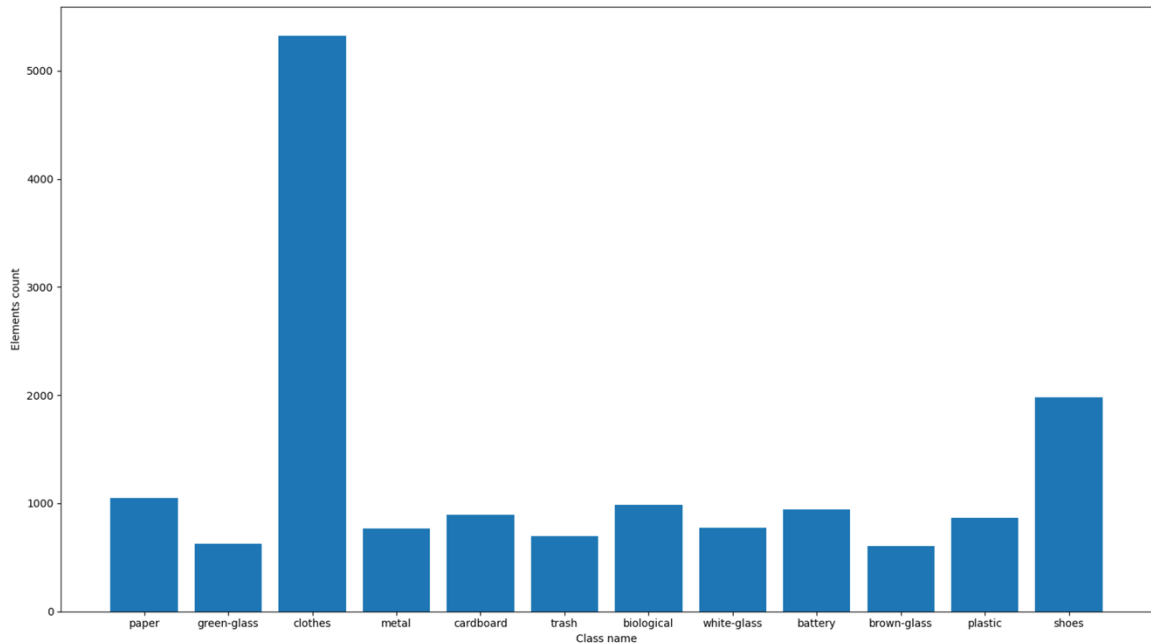


Fig. 1 Dataset distribution

As a base model MobileNetV2 1.0 was taken. The decision was made due to the number of parameters, which is equal to 3.50M. It makes model work much faster and takes less space comparing to ResNet-50 with 25.55M parameters and ResNet-34 with 21.79M parameters. As a classification mechanism, fully connected layer was placed on top of MobileNetV2 receiving 1280 and producing 12 out features. Loss function which is applied to this architecture is Cross Entropy and computes loss between target and input. In PyTorch implementation of CrossEntropyLoss logits are directly passed to loss function, so there is no need to apply SoftMax before computing loss while SoftMax and logarithm are going to be applied automatically. This results into receiving logits as an output of the model with shape (batch size, 12). To calculate one predicted class argmax function was used, which returns index of the biggest logit, meaning that model is more likely to predict that returned class index. Optimizer used during training is Adam with learning rate equals to 0.1. Learning rate was changed during training process in order not to stuck in local function minimum. As a metrics to prove that model is learning was accuracy taken [6]. Accuracy is computed as correct predictions divided by all predictions.

It's seen a class imbalance to clothes class due to big number of images in this category, but it can be resolved by using under-sampling, over-sampling, or using a class balance technique. Over-sampling was not used, while there is not enough data to balance all classes, under-sampling may lead towards overfitting data, due to dataset size decrease. In addition to a fact, that we use smaller neural net comparing to ResNet-50, we selected RandomWeightedSampler that is available in PyTorch framework. By assigning a weight to each sample in data, we get nearly equal amount of all classes passed to a neural network during training and don't lose any image from dataset, which prevents overfitting to certain types of images.

IV. RESULTS

Training neural net on basic dataset produced accuracy equals 0.9192 on validation set and 0.8713 on test dataset, but deeper analysis using confusion matrix showed that neural net overfitted towards clothes class. After applying RandomWeightedSampler with data augmentations result was improved. As a result, neural network described in this paper shows validation accuracy 86.6537 and validation loss equals to 1.0779. Learning process can be seen on fig.2 graphs.

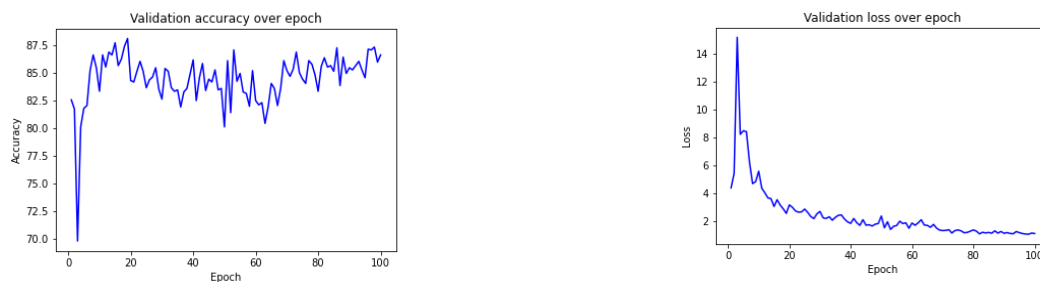


Fig. 2. Validation loss and validation accuracy per epoch

When analyzing results, one more confusion matrix was built, which showed that data balancing and data augmentations helped neural network prevent overfitting towards clothes class. High number of wrong classifications occurs only between classes white-glass and plastic. On normalized confusion matrix it results into score equals to 0.12, when right classification for white-glass shows 0.59 and 0.61 for plastic.

V. DISCUSSION

Using ResNet in this classification problem would give better results, but also decrease performance of the model. Using MobileNet as a base model gives us a good starting point with understanding of whether it's possible to solve multiclass classification task using smaller models. In my opinion, better results can be produced if experimenting more with training data, instead of model architecture. All types of glass can be combined into one class, cardboard and paper can also be merged. This may also resolve the issue with wrong classifications of white-glass and plastic. Overall, the model should be retrained for exact task, because sorting all types of garbage using MobileNet may not be as effective, but it may be good enough to classify 4 garbage classes, for example.

VI. FUTURE RESEARCH

As future research, I believe that gathering more data for smaller classes, like paper and plastic may lead to better model performance. Also, some more data augmentations may improve the model accuracy, while dataset was gathered from different resources, which resulted to images taken in different lightning conditions. To produce better results, model should be trained with augmentations, that are similar to environment where the model is going to be used.

VII. CONCLUSION

As a result, a neural network was built that can classify 12 types of household waste with accuracy 86.6537. This result was achieved with the help of transfer learning technique. MobileNetV2 1.0 was taken as a base model, resulting in 3.50M parameters which makes this neural network lightweight and possible to run in real-time on embedded devices. When training on base dataset neural network shows higher accuracy but based on confusion matrix, we conclude that this happened due to overfitting to clothes class. Applying class balance techniques and data augmentations resulted into model performance well on all classes, except plastic and white paper.

VIII. ACKNOWLEDGMENT

The author gratefully acknowledges the contributions of professors at Odesa Polytechnic National University, for assistance in obtaining the knowledge necessary to conduct such research.

IX. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] Oecd, Global Plastics Outlook. 2022, σ. 201.
- [2] Z. Kang, J. Yang, G. Li, και Z. Zhang, 'An Automatic Garbage Classification System Based on Deep Learning', IEEE Access, τ. PP, σσ. 1–1, 07 2020.
- [3] M. Mohamed, "Garbage classification (12 classes)," Kaggle, 24-Jan-2021. [Online]. Available: <https://www.kaggle.com/datasets/mostafaabla/garbage-classification>. [Accessed: 30-Oct-2022].
- [4] Cchanges, "Garbage classification," Kaggle, 24-Nov-2018. [Online]. Available: <https://www.kaggle.com/datasets/asdasdasdas/garbage-classification>. [Accessed: 30-Oct-2022].
- [5] Ololo, "Clothing dataset (full, high resolution)," Kaggle, 19-Oct-2020. [Online]. Available: <https://www.kaggle.com/datasets/agrigorev/clothing-dataset-full>. [Accessed: 30-Oct-2022].
- [6] M. Grandini, E. Bagli, και G. Visani, 'Metrics for Multi-Class Classification: an Overview'. arXiv, 2020.

Моделювання імпедансних спектрів, отриманих для системи пористий електрод/електроліт

Наталія Іванічок¹, Олег Іванічок¹, Софія-Вікторія Склепова¹,
Павло Колковський², Богдан Рачій¹, Роман Лісовський³

¹Кафедра матеріалознавства і новітніх технологій,
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна,

²Відділ хімії твердого тіла,
Інститут загальної та неорганічної хімії імені В. І. Вернадського НАН України,
Київ, Україна.

³Кафедра медичної інформатики, медичної і біологічної фізики,
Івано-Франківський національний медичний університет,
Івано-Франківськ, Україна

Змінюючи температуру карбонізації рослинної біомаси (шкарлупи волоських горіхів), отримано пористі вуглецеві матеріали (ПВМ) з різним розподілом пор за розміром та великою питомою площею поверхні. Методом імпедансної спектроскопії досліджено та промодельовані фізико-хімічні процеси, що відбуваються на межі розділу пористий вуглецевий електрод/електроліт.

Ключові слова—Пористий вуглецевий матеріал, Імпедансна спектроскопія, Питома ємність.

I. ВСТУП

Швидкість процесів накопичення заряду за рахунок формування подвійного електричного шару є одним з ключових факторів, який головним чином залежить від сумісності розміру йонів електроліту з розподілом пор за розміром матеріалу [1]. Щоб зрозуміти фундаментальний зв'язок між структурними характеристиками та електрохімічними властивостями (зокрема, питомою ємністю), необхідний комплекс електрохімічних досліджень. Для дослідження фізико-хімічних процесів, що відбуваються на межі розділу вуглецевий електрод/електроліт, був використаний метод спектроскопії електрохімічного імпедансу.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Отримання нових або модифікація відомих електродних матеріалів для суперконденсаторів є необхідним методом для досягнення високих електроємнісних характеристик як матеріалу, так і електрохімічної системи в цілому [2, 3]. Для поліпшення енергоємнісних характеристик електрохімічних конденсаторів використовуються численні стратегії, пов'язані зі зміною структури та збільшенням площі поверхні матеріалів для електродів [4-6]. Властивості електрохімічних систем зручно вивчати за відгуком цієї системи на зовнішній синусоїдальний сигнал. У цьому плані найбільш інформативним є метод імпедансної спектроскопії, оскільки він дає можливість проводити дослідження в досить широкому інтервалі частот ($10^6 - 10^{-3}$ Гц) [7]. Аналіз поведінки систем електрод/електроліт здійснюється шляхом моделювання імпедансних спектрів (діаграм Найквіста) еквівалентною електричною схемою.

III. МЕТОДИ

ПВМ отримали методом термічної карбонізації при температурі 800 °С вихідної сировини природного походження. В якості електроліту використовували 33 % водний розчин солі КОН.

Електрохімічні дослідження ПВМ проводилися на спектрометрі Autolab PGSTAT/FRA-2 (Нідерланди), використовуючи методи імпедансної спектроскопії.

При проходженні змінного струму згідно закону Ома повний комплексний опір (імпеданс) Z^* визначається виразом [8]:

$$Z^* = \frac{E_0 e^{j\omega\tau}}{I_0 e^{j\omega\tau + \varphi}} = \frac{E_0}{I_0} e^{-j\varphi} = Z \cos \varphi - jz \sin \varphi = z' - jz'',$$

де E_0 , I_0 – амплітуди напруги і струму, ω – кутова частота, φ – кут зсуву фаз, j – уявна одиниця ($j = \sqrt{-1}$), Z' , Z'' – дійсна (активна), і уявна (реактивна) складові імпедансу. На практиці графічна залежність Z^* в координатах Z' , Z'' носить назву годограф імпедансу або його спектр.

Основна задача дослідження зводиться до побудови еквівалентної схеми (моделі), що адекватно відображає електрофізичні і хімічні процеси в зразку, а також до розрахунку різних параметрів цієї моделі. Модель відображає основні властивості досліджуваного зразка, служить для передбачення його поведінки в різних умовах.

IV. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ

Діаграми Найквіста, отримані для електрохімічної системи електрод/електроліт в діапазоні частот $10^{-2} \div 10^5$ Гц, представлено на рис. 1. В діапазоні високих частот спостерігається невелика ділянка у вигляді півкола (вставка на рис. 1), що свідчить про наявність незначних фарадеївських процесів на межі розділу вуглецевий електрод/електроліт [9]. Одразу після півкола можна спостерігати похилу лінійну ділянку, що є притаманною поведінкою для вуглецевого електрода. Це може бути пов'язано з процесом проникнення електроліту у пори різного розміру [10].

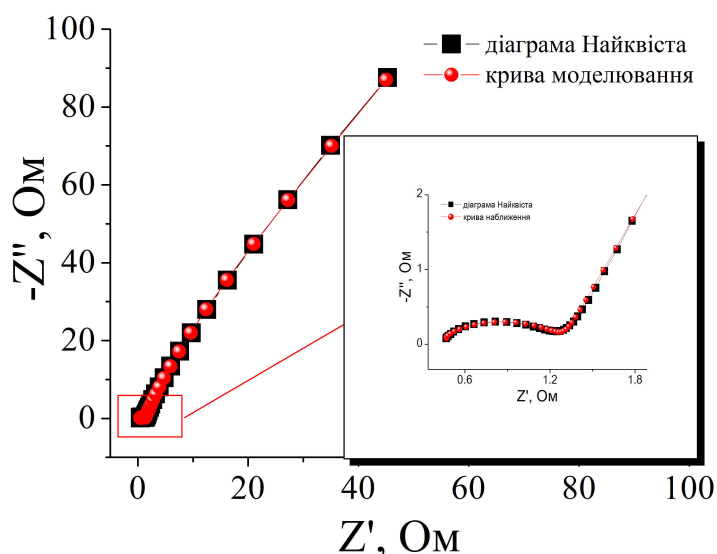


Рис. 1. Діаграма Найквіста (отримана при напрузі відкритого кола) та крива моделювання для системи електрод/електроліт

Оптимальні результати моделювання експериментальних кривих $Z'(-Z'')$ (рис.1) отримані при застосуванні електричної еквівалентної схеми (ЕЕС), яка представлена на рис. 2. ЕЕС включає в себе послідовний еквівалентний опір R_1 (опір електроліту, контактів та підвідних проводів), ланки $CPE_1 \parallel R_2$, яка відповідає за дифузійні процеси в макропорах вуглецевого матеріалу та накопичення заряду на межі розділу вуглецевий матеріал/електроліт, та $C_1 \parallel R_3-CPE_2$ – відображає вказані процеси в мезо- і мікропорах. Елементи постійної фази CPE використано для забезпечення гнучкості моделювання. CPE_1 – елемент постійної фази ємнісного типу, який відображають процеси накопичення заряду в макропорах вуглецевого матеріалу, CPE_2 – елемент постійної фази дифузійного характеру, що пов'язаний з процесами обмеженої дифузії йонів K^+ в мікропорах, R_2 і R_3 – опори перенесення заряду в вуглецевому матеріалі.

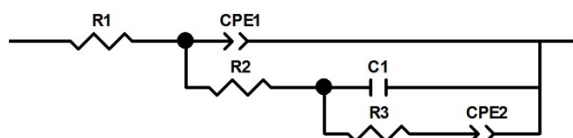


Рис. 2. ЕЕС, використана для апроксимації імпедансних спектрів

Запропонована ЕЕС (рис. 2) добре підходить для наближення експериментальних імпедансних спектрів в широкому діапазоні частот, а квадрат відхилення кривої наближення від експериментального спектру становить 3×10^{-4} . Використання ЕЕС дала можливість отримати значення параметрів, що входять в схему (табл. 1).

Таблиця 1

Параметри ЕЕС для електрохімічної системи електрод/електроліт

Напруга, В	R_1 , Ом	CPE_{1T} , мФ	CPE_{1P}	R_2 , Ом	C_1 , Ф	R_3 , Ом	CPE_{2T} , Ф	CPE_{2P}
-0,2	0,41	1,9	0,72	0,95	0,03	0,55	0,05	0,67
-1	0,55	4,8	0,63	0,77	0,19	0,60	0,98	0,79

Оскільки при низьких частотах активний опір практично не залежить від частоти, то можна оцінити значення ємності досліджуваного ПВМ, яка буде залежати від прикладеної напруги. Згідно співвідношення $C = 1/2\pi\nu Z''$ [8] та врахувавши масу електрода, було обчислено питому ємність ПВМ. При напрузі відкритого кола питома ємність дослідженого ПВМ становить 3,3 Ф/г, а при напрузі зміщення -1 В становить 84 Ф/г.

V. ВИСНОВКИ

Методом імпедансної спектроскопії отримано спектри та змодельовано фізико-хімічні процеси, що відбуваються на межі розділу пористий електрод/електроліт. Запропоновано еквівалентну електричну схему для моделювання експериментальних імпедансних спектрів. Квадрат відхилення змодельованої кривої від експериментального спектру становить 3×10^{-4} . Використання ЕЕС дало можливість отримати значення деяких фізичних параметрів, які характеризують електрохімічну систему електрод/електроліт. Обчислено питому ємність ПВМ, яка при напрузі відкритого кола становить 3,3 Ф/г, а при напрузі зміщення -1 В – 84 Ф/г.

ЛІТЕРАТУРА

- [1] B. Dyatkin, N.C. Osti, A. Gallegos, Y. Zhang, E. Mamontov, P.T. Cummings, J. Wu, Y. Gogotsi, "Electrolyte cation length influences electroadsorption and dynamics in porous carbon supercapacitors", *Electrochim. Acta* vol. 283, pp. 882–893, 2018

- [2] Ivanichok, N., Budzuliak, I., Moiseienko, M., Lisovskiy, R., Rachii, B., Gamarnyk, A., Turovska, L., & Lisovska, S. "Electrochemical properties of nanoporous carbon materials obtained from raw materials of plant origin (hemp shives) ", *Physics and Chemistry of Solid State*, vol. 21, no. 1, pp. 35–42, 2020
- [3] W.L. Zhang, J.H. Xu, D.X. Hou, J. Yin, D.B. Liu, Y.P. He, H.B. Lin, "Hierarchical porous carbon prepared from biomass through a facile method for supercapacitor applications", *J. Colloid Interface Sci.*, vol. 530, pp. 338–344, 2018
- [4] I.F. Myronyuk, V.I. Mandzyuk, V.M. Sachko, R.P. Lisovsky, B.I. Rachiy, "Morphological and electrochemical properties of the lactose-derived carbon electrode materials", *Journal of Nano- and Electronic Physics*, vol. 8, no. 4, P. 04006, 2016
- [5] M. Liu, F. Wei, X. Yang, S. Dong, Y. Li, X. He, "Synthesis of porous graphene-like carbon materials for high-performance supercapacitors from petroleum pitch using nano-CaCO₃ as a template", *Xinxing Tan Cailiao/New Carbon Mater.*, vol. 33, pp. 316–323, 2018
- [6] V.I. Mandzyuk, I.F. Myronyuk, V.M. Sachko, I.M. Mykytyn, "Template synthesis of mesoporous carbon materials for electrochemical capacitors", *Surface Engineering and Applied Electrochemistry*, vol. 56, no. 1, pp. 93–99, 2020
- [7] V.I. Mandzyuk, I.F. Myronyuk, N.Ya. Ivanichok, B.I. Rachiy, "Impedance spectroscopy of capacitor systems based on saccharide-derived porous carbon materials", *Physics and Chemistry of Solid State*, vol. 22, no. 4, pp.711–716, 2021
- [8] З.Б. Стойнов, Б.М. Графов, Б. Савова-Стойнова, В.В. Елкин. *Электрохимический импеданс*, Москва: Наука, 1991, 331 с.
- [9] Tonurist, K.; Vaas, I.; Thomberg, T.; Jänes, A.; Kurig, H.; Romann, T.; Lust, E. "Application of multistep electrospinning method for preparation of electrical double-layer capacitor half-cells", *Electrochim. Acta*, vol. 119, pp. 72–77, 2014
- [10] L.I. Jing and Zhao Xiao-dong. "Preparation and Electrochemical Characterization of an Activated Carbon Material of High Surface Area for Supercapacitor", *Advanced Materials Research*, vols 463–464, pp. 410–414, 2012

Класифікація структурних елементів програмного забезпечення у задачах генерації тестів

Роман Базилевич, Франко Андрій

*Кафедра програмного забезпечення
Національний університет “Львівська Політехніка”
Львів, Україна*

Анотація—Проаналізовано методи класифікації функцій, як структурних елементів програмного забезпечення у задачах генерації модульних тестів. Визначено основні переваги та недоліки наявних методів класифікації функцій. Запропоновано комплексний алгоритм класифікації функцій для задач генерації тестів. Описано його переваги над наявними методами.

Ключові слова— *Якість програмного забезпечення, символічне виконання, абстрактне синтаксичне дерево, тестування методом білої скриньки*

I. ВСТУП

Розв’язання задачі автоматизованого генерування тестів є важливим для підвищення швидкості розробки програмного забезпечення. Основним завданням генерування тестів є визначення очікуваних результатів тесту та вхідних даних для нього. Вирішення даного завдання вимагає виконання класифікації та розпізнавання структурних елементів програмного забезпечення. Це дасть змогу визначити їх призначення та згенерувати відповідні тести.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Основним структурним елементом програмного забезпечення, що використовується для генерації тестів є функція. Для класифікації функцій найпоширенішим методом є аналіз абстрактного синтаксичного дерева [1-2]. Існують системи засновані на семантичному аналізі [1,3]. Основним засобом генерування тестів є символічне виконання. Воно використовує евристичні методи для генерації вхідних даних на основі програмного коду функції [4]. Для удосконалення символічного виконання необхідно мати методи отримання додаткових метаданих про функцію. Це дасть змогу генерувати вхідні дані, що відповідають реальному використанню функції. Класифікація функцій дозволить згенерувати додаткові метадані, що будуть відображати причино-наслідкові зв’язки між очікуваним результатом роботи функції та вхідними даними для тесту. Наявність подібних метаданих створить можливості для накопичення моделей для тестування стандартних функцій (перевірка контрольної суми пакета даних, взаємодія зі стандартними структурами даних, використання поширених патернів).

III. МЕТОДИ

Для створення механізму розпізнавання функцій у процесі тестування програмного забезпечення необхідно визначити: поняття ідентичності двох функцій, методи порівняння функцій та характеристик для порівняння. Характеристики для порівняння функцій у рамках задачі генерування тестів були визначені у попередньому дослідженні [5]. На основі визначених характеристик функції буде проведена оцінка наявних методів ідентифікації

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

функції та запропоновано новий метод. Для моделювання роботи наявних методів та формування нової функції F визначено як набір інструкцій i_1-i_n , де n це їх кількість, які оперують параметрами p_1-p_k (множина P), де k є кількістю параметрів. В результаті виконання інструкцій i_1-i_n (множина I), можуть зазнавати зміни або зчитуватися дані за межами локальної області видимості функції d_1-d_m (множина D), де m це їх кількість. В результаті послідовного виконання інструкцій утворюється множина даних D_I (модифікована множина D) та результат роботи функції O . Для кожної функції існує залежність між множиною вхідних даних (об'єднання множин P та D) та множиною усіх вихідних даних (об'єднання O з D_I). В залежності від значень елементів множин P та D формується значення D_I та O . Ідентичність функцій означає, що для всіх можливих значень вхідних параметрів p_1-p_k та даних поза локальною областю видимості d_1-d_m обидві функції продукують ідентичні D_I та O . Повна ідентичність означає також ідентичність усіх інструкцій, що є умовами або циклами. Формула (1) демонструє описані вище відношення між входом та виходом функції.

$$F(P, D) = D_I \cup O \quad (1)$$

IV. РЕЗУЛЬТАТИ

Першою групою методів є порівняння інструкцій i_1-i_n для двох функцій. Порівняння інструкцій дає змогу з певною імовірністю оцінити подібність функцій [1-2]. Для задачі генерації тестів такої подібності не достатньо. Функції можуть мати різний набір інструкцій та бути ідентичними для задачі генерації тестів. Синтаксис мов програмування дозволяє створювати множину конструкцій, що виконують однакову задачу, але різняться між собою. Прикладом є математичні обчислення, що можуть мати різні реалізації, але приводити до ідентичного результату. В цьому випадку множини I_1 та I_2 мають суттєві між елементами i_1-i_n . Тести згенеровані для них будуть ідентичними та міститимуть однакові вхідні дані p_1-p_k та зовнішні дані d_1-d_m . Другою групою методів є семантичний аналіз [3], аналіз назв змінних та коментарів. Він не є надійними через різні стилі програмування та вимоги обробки людської мови. Відсутність зв'язку з характеристиками функції робить семантичний аналіз лише додатковим засобом.

Для розпізнавання функції в задачі тестування необхідно встановити відповідність між вхідним та вихідними даними та порівняти утворені залежності за формулою (1) з аналогічними даними для інших функцій. Отримання цих даних вимагає використання символного виконання та значного обчислювального часу [4]. Збіг залежностей вихідних даних від вхідних не гарантує ідентичність тестових наборів даних. Дані, які генерують евристичні методи під час символного виконання залежать від набору інструкцій i_1-i_n . У разі відмінності в умовних операторах може виникати різниця в обрахованих наборах вхідних даних для тестування. Тому запропоновано комплексний алгоритм, що враховує структуру функцій та відношення між вхідним та вихідними даними. Його подано на рис. 1. Він дає змогу оцінювати подібність багатьох функцій між собою внаслідок економії обчислювальних ресурсів. На першому етапі виконується лінійне порівняння абстрактних синтаксичних дерев. На другому запропоновано використовувати методику символного виконання для фінальної оцінки подібності. Порівняння абстрактних синтаксичних дерев пропонується здійснювати з використанням вже розроблених методів для перевірки на плагіат [1-3]. Символьне виконання буде використано з метою побудови таблиць, що міститимуть класи вхідних даних та відповідні їм результати роботи функції. Отримані таблиці будуть порівнянні між собою для обраних функцій. Це гарантує структурну та функціональну подібність. Поріг s буде використано для встановлення мінімального допустимого відхилення при класифікації функцій.

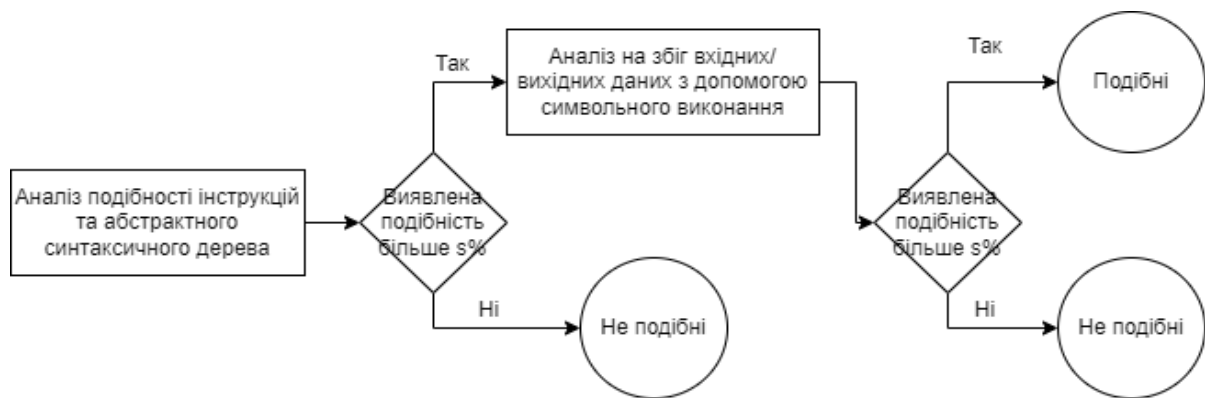


Рис. 1. Запропонований алгоритм класифікації функції для задач генерації тестів

V. ОБГОВОРЕННЯ

Застосування такого алгоритму дозволить проводити розпізнавання функцій при генеруванні тестів. Можливості класифікації функцій дадуть змогу порівнювати вихідний код з базою даних стандартизованих функцій. Це дасть змогу отримати додаткові відомості про призначення функцій та генерувати на їх основі додаткові тести.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Результати даного дослідження надалі можуть бути застосовані для удосконалення систем символічного виконання та генерування тестів. Розпізнавання стандартизованих функцій дасть змогу генерувати тести, що тестують розпізнану функцію на виконання її призначення, яке було визначено з допомогою класифікації.

VII. ВИСНОВКИ

Проаналізовано проблему класифікації функції у задачах генерації модульних тестів. Запропоновано використовувати комплексний алгоритм класифікації з використанням порівняння абстрактних синтаксичних дерев та символічного виконання. Вказано на необхідність враховувати залежність між вхідними та вихідним даним функції та її структуру при класифікації функцій для задачі генерування тестів. Описано можливості покращення систем генерування тестів з допомогою запропонованого підходу.

ЛІТЕРАТУРА

- [1] S. Yang, L. Cheng, Y. Zeng, Z. Lang, H. Zhu and Z. Shi, "Astera: Deep Learning-based AST-Encoding for Cross-platform Binary Code Similarity Detection," *2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 2021, pp. 224-236.
- [2] T. Shippey, D. Bowes and T. Hall, "Automatically identifying code features for software defect prediction: Using astN-grams", *Inf. Softw. Technol.*, vol. 106, pp. 142-160, 2019.
- [3] A. Mahmoud and G. Bradshaw, "Semantic topic models for source code analysis", *Empirical Softw. Eng.*, vol. 22, no. 4, pp. 1965-2000, Aug. 2017.
- [4] C. Cadar and M. Nowack, "KLEE symbolic execution engine in 2019", *Int. J. Softw. Tools Technol. Transf.*, vol. 23, no. 6, pp. 867-870, 2021.
- [5] R. Bazylevych and A. Franko, "Hierarchical model of automated test generation system", *Scientific Bulletin of UNFU*, vol. 32(4), pp. 77-83, Aug. 2022.

Пристрій для дослідження поверхневого натягу рідин методом газового струменя

Роман Боднар

*Кафедра метрології та інформаційно-виміральної техніки
Івано-Франківський національний технічний університет нафти і газу
Івано-Франківськ, Україна*

Анотація – Розроблено пристрій для дослідження залежності геометричних параметрів меніска рідини від поверхневого натягу під дією газового струменя.

Ключові слова – Поверхневий натяг, Вимірювання, Поверхнево-активні речовини, Меніск.

I. ВСТУП

У багатьох технологічних процесах флотації руд, нафтогазовидобутку, буріння свердловин, для очищення виробів на виробництві і в побуті, а також при виробництві штучних волокон, каучуку, пластмас, в целюлозно-паперовій та інших галузях використовують розчини поверхнево-активних речовин (ПАР). З метою оптимального проведення таких процесів проводиться постійний або періодичний моніторинг поверхневого натягу (ПН) технологічних рідин для оцінки вмісту в рідині ПАР [1].

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Принцип роботи багатьох сучасних приладів для визначення ПН рідин базується на таких методах як лежачої краплі, відриву кільця, максимального тиску в бульбашці та ін.[2]. Кожен з цих методів має конкретне застосування в залежності від виду рідин та умов вимірювання ПН, але в усіх цих методах є одна спільна вада – контакт рідини з діючим на неї інструментом впливу. Таким інструментом є платинове кільце у методі відриву кільця, калібрований капіляр у методі максимального тиску в бульбашці чи пластина у методі Вільгельмі і т. п. Наявність будь-якого інструмента впливу вимагає його ретельного очищення після кожного вимірювання, що приводить до зниження продуктивності процесу визначення ПН рідин та унеможливає проведення постійного моніторингу ПН. Недотримання чистоти інструменту впливу приведе до значних похибок в результатах вимірювання.

Поверхневий натяг є реальною силою, яка діє вздовж поверхні рідини і при розтягуванні поверхні вона виконуватиме роботу

$$A = \sigma \cdot dS, \quad (1)$$

де σ – поверхневий натяг; S – площа поверхні рідини.

Робота деформації реального поверхневого шару рідини представляється як сума робіт деформації деякої геометричної поверхні, яка має ПН σ , і неминучої супутньої деформації об'ємних фаз. У загальному випадку робота деформації елементів поверхні включає члени, які зв'язані із змінами об'єму, викривленням поверхні і зміною положення в гравітаційному полі.

Локальні тиски P_1 та P_2 беруться по обидві сторони поверхні, яка розділює рідину і газову фазу з врахуванням сили тяжіння. В результаті таких міркувань отримується рівняння [2]:

$$P_1 - P_2 = \sigma \left(\frac{1}{R_1^2} + \frac{1}{R_2^2} \right) - \left(\frac{C_1}{R_1^2} + \frac{C_2}{R_2^2} \right) + mg \cos \varphi, \quad (2)$$

де R_1, R_2 – радіуси кривизни поверхні у взаємно перпендикулярних площинах; C_1, C_2 – коефіцієнти, зв'язані з деформацією відповідних елементів поверхні; m – маса елемента поверхні; g – прискорення сили земного тяжіння; φ – кут між вертикаллю і нормаллю до поверхні.

III. МЕТОДИ

Рівняння (2) виражає умову рівноваги між рідкою і газовою фазами, які розділяють викривлену поверхню, і з нього можна визначати ПН σ незалежно від способу створення деформації поверхні розділу фаз.

Для визначення ПН рівняння (2) реалізується на практиці, зокрема у методі максимального тиску в бульбашці, який базується на класичному рівнянні Лапласа, що отримується із (2) при певних допущеннях.. Але капілярна поверхня утворюватиметься не тільки всередині бульбашки, що витискується із капіляра в рідину, але й при дії газового струменя на вільну поверхню рідини. При цьому геометричні параметри меніска будуть пов'язані з ПН рідини. Для встановлення такого зв'язку розроблено пристрій, структурна схема якого показана на рис.1

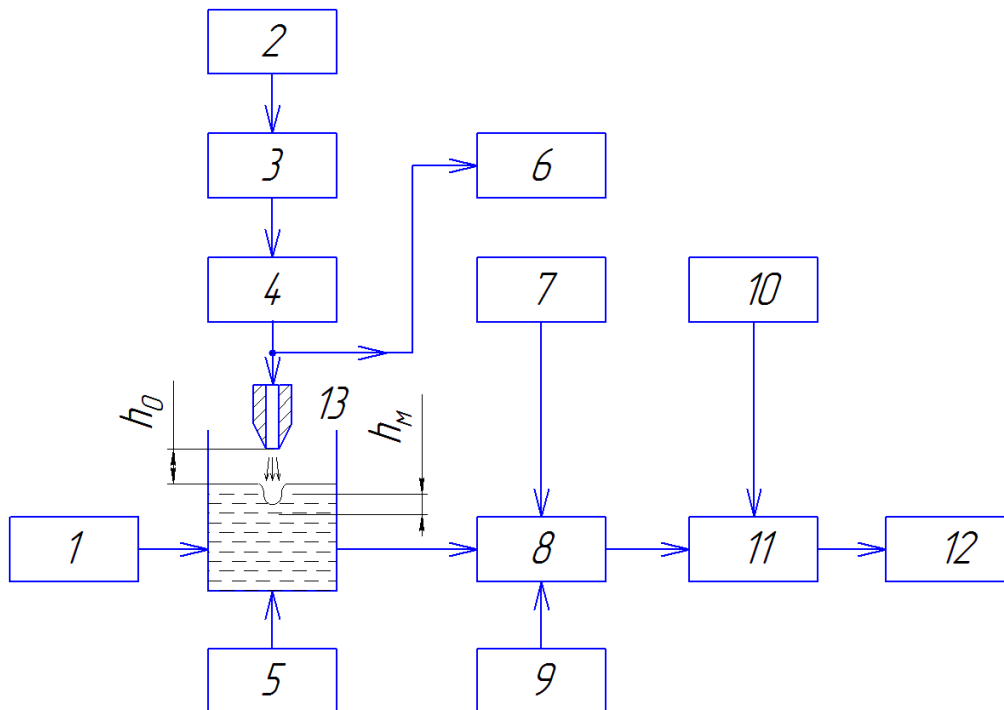


Рис.1 – 1 – освітлювач; 2 – мікрокомпресор; 3 – стабілізатор тиску; 4 – регулятор тиску; 5 – підймальний механізм; 6 – вимірювач тиску; 7 – регулятор наведення чіткості зображення; 8 – цифровий мікроскоп; 9 – механізм переміщення; 10 – клавіатура; 11 – персональний комп'ютер; 12 – монітор; 13 – сопло

Робота пристрою полягає в наступному. Повітря під тиском із мікрокомпресора 2 поступає на стабілізатор 3 і регулятор тиску 4, до якого приєднано вимірювач тиску 6 і сопло 13, з якого повітря видувається на поверхню рідини. Під дією тиску повітря на поверхні рідини створюється меніск, форму якого спостерігають за допомогою цифрового мікроскопа 8. Освітлення контура меніска створюється освітлювачем 1. Інформація з

цифрового мікроскопа поступає на персональний комп'ютер 11, який буде обробляти цю інформацію.

IV. РЕЗУЛЬТАТИ

Результати перших досліджень на розробленім пристрою показано на рис. 2–3.

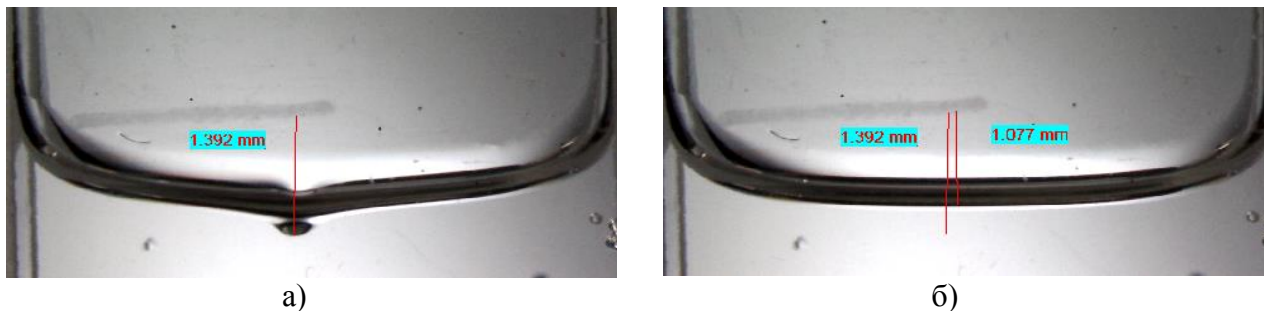


Рис.2 – Зображення профілів менісків води під дією повітряного струменя (а) і в спокійному стані (б)

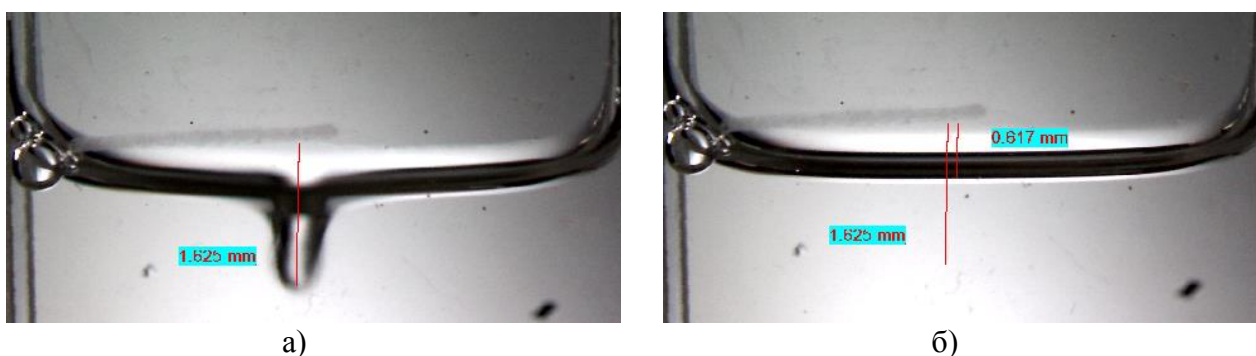


Рис.3 – Зображення профілів менісків розчину ПАР під дією повітряного струменя (а) і в спокійному стані (б)

Під дією повітряного струменя на поверхні води утворюється воронка (рис. 2, а), глибина якої дорівнює:

$$h_b = l_1 - l_2, \quad (3)$$

де l_1 – віддаль від вибраної реперної лінії до вершини меніска; l_2 – віддаль від вибраної реперної лінії до поверхні води в спокійному стані.

Віддалі l_1 і l_2 визначались за допомогою цифрового мікроскопа (рис. 2, б). Для води глибина воронки становила $h_b = 0,315$ мм.

Після додавання краплі ПАР у кювету з водою при незмінних інших параметрах експерименту утворюється воронка (рис. 3, а), глибина якої становила $h_p = 0,808$ мм.

V. ОБГОВОРЕННЯ

Отримані результати показують очевидну залежність геометричних параметрів меніска рідин під дією повітряного струменя від величини ПН рідин. Але початкові результати показують, що ця залежність є нелінійною. Використання запропонованого методу визначення ПН дає можливість здійснювати вимірювання без затрат праці і часу на промивання вимірювальних інструментів. Це дозволяє проводити неперервний моніторинг ПН рідин в технологічних процесах із застосуванням ПАР, а також екологічний контроль стічних вод.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальші дослідження повинні бути спрямовані на встановлення точного зв'язку геометричних параметрів меніска рідин під дією повітряного струменя від величини ПН рідин і метрологічної оцінки запропонованого методу.

VII. ВИСНОВКИ

Результати перших досліджень підтверджують припущення про можливість використання рівняння (2) для визначення ПН неконтактним способом. Геометричні зміни профілю меніска під дією повітряного струменя при зміні ПН рідини існують, але зв'язок між ними є складний, оскільки коефіцієнти C_1 , C_2 в рівнянні (2) визначити досить складно, інші ж величини в принципі визначаються.

ЛІТЕРАТУРА

1. В. Д Михайлюк., М. І. Рудий, Р.Т. Боднар та ін., “Використання поверхнево-активних речовин в процесах нафтовидобутку на родовищах ВАТ “Укрнафта”. Івано-Франківськ, Україна: ПП “Галицька друкарня ПЛЮС”, 2009, с. 400
2. А.И. Русанов, В.А. Прохоров “Межфазная тензометрия”. Санкт-Петербург, РФ: Химия, 1994, с. 400

Development of forestry application using deep learning technology

Nataliia Liubchenko, Andrii Podorozhniak, Daniil Onishchenko

*Department of Informatics and Information Property
National Technical University "Kharkiv Polytechnic Institute"
Kharkiv, Ukraine*

Abstract — Forestry – one of the fields of human activity, which has one main problem impossibility to keep the records. The main goal of the work is to create application, which will make analyse of forests with large area in time possible with help of deep learning methods. The work presents many methods of remote sensing of the Earth, index images, as well as comparison of modern neural networks architectures for general image analyses for forestry.

Keywords — *Forests, Remote Sensing of the Earth, Index Images, Artificial Intelligence, Deep Learning, Artificial Neural Networks.*

I. INTRODUCTION

Forestry is the science and craft of creating, managing, planting, using, conserving and repairing forests, woodlands, and associated resources for human and environmental benefits [1]. The main problem of the forestry is impossibility to keep the records of the forest in terms of large area [2]. Modern technologies of artificial intelligence allow creation of new information systems for automation of work in all of human fields. One of such fields is forestry, whose one of the characteristics is largeness, because of large area of analysed territory. Because of remote sensing of the Earth we can now obtain images of earth's territory in different timestamps, that allows the analysis of changes in the vegetation.

II. RELATED WORK

The research of satellite images by the information systems is popular topic for researches, as it allows them to tests a large number of different ideas on real-life tasks [3, 4]. One type of such systems is classification systems, which can determine the change in the vegetation [5]. For instances, one of the study [6] showed the dependency of change in vegetation with change of index images. The idea of this work is to create a universal system for green cover changes recognition on the territory of Ukraine.

III. METHODOLOGY

Remote sensing allows obtaining high-precision images with the help of NASA's Landsat 8 satellites and the European Space Agency's Sentinel-2. These make possible to produce large in scale studies and use the system as a universal for many types of land cover (forests, fields, mountains, etc.). Unfortunately, it's almost impossible to process images directly, since satellites save data in the form of many channel images with different wavelengths. Therefore, it is absolutely necessary to use index images, which are created from a combination of channels [7]. The most popular type of such images is NDVI (Normalized difference vegetation index) [8], because the main advantage of these images is general-purpose analysis of green cover, as NDVI doesn't influenced so much by noise. This index is created from only two channels – red and infrared. Next step is neural network processing of index image. Because of large number of different neural network architectures nowadays, it makes sense to conduct an analysis to determine which one will work best.

IV. RESULTS

As the result comparative characteristics of modern neural network's architectures were presented. Minimum Top 1 accuracy among all the model types is around 80% was achieved. In addition, the artificial intelligence model were created to determine a change of the vegetation on the territory of Ukraine.

V. FUTURE RESEARCH

This topic makes future research possible in many different ways. For example, the more complex combination of index images as an input to the neural network would show better results. One of possible way of improvement is usage of recurrent neural networks. There are not so many studies on this topic, due to the complexity of analysis. However, the results of such research will be a significant contribution to the image processing development.

VI. CONCLUSION

Created information system for forestry record keeping is the main purpose of the work. As the result of analysis and synthesis of necessary methods and technologies geoinformation system was created to determine vegetation changes on the territory of Ukraine.

REFERENCES

- [1] D. Vollmuth, "The changing perception of coppice with standards in German forestry literature up to the present day – From a universal solution to a defamed and overcome evil – and back?", *Trees, Forests and People*, 10, 2022, 100336. DOI: 10.1016/j.tfp.2022.100338
- [2] E. Hrynyk and O. Vasylyuk, "Self-sowed forest, or a simple way to increase the number of forests of Ukraine," Ukrainian environmental group, 16-Oct-2019. [Online]. Available: <https://uncg.org.ua/samosijni-lisy/>. [Accessed: 30-Oct-2022].
- [3] V. Yaloveha, A. Podorozhniak, and H. Kuchuk, "Convolutional neural network hyperparameter optimization applied to land cover classification", *Radioelectronic and Computer Systems*, n. 1, 2022, pp. 115-128. DOI: 10.32620/reks.2022.1.09.
- [4] A. Podorozhniak, N. Liubchenko, M. Kvochka, and I. Suarez, "Usage of intelligent methods for multispectral data processing in the field of environmental monitoring", *Advanced Information Systems*, vol. 5, n. 3, 2021, pp. 97-102. DOI: 10.20998/2522-9052.2021.3.13
- [5] D. Onishchenko, N. Liubchenko, and A. Podorozhniak, "Vykorystannia tekhnolohii hlybokoho navchannia dlia rozpoznavannia molodykh samosiviv (Using deep learning technologies to recognize young self-sowed forest)", Proceedings of the 9th international scientific and technical conference Problems of informatization, Ukraine, 2021, vol. 2, p. 66 [Online]. Available: http://repository.kpi.kharkov.ua/bitstream/KhPI-Press/58385/1/Onishchenko_Vykorystannia_2021.pdf. [Accessed: 30-Oct-2022]
- [6] D. R. Luna, J. M. Manzano, B. M. Bonilla, and J. H. Garcia, "Analysis of the Vegetation Indices NDVI, GNDVI, and NDRE for the Characterization of Coffee Crops", *Ingeniería y Desarrollo*, vol. 38, n. 2, 2020, pp. 298-312. DOI: 10.14482/inde.38.2.628
- [7] V. Yaloveha, D. Hlavcheva, and A. Podorozhniak, "Spectral Indexes Evaluation for Satellite Images Classification using CNN", *Journal of Information and Organizational Sciences (JIOS)*, vol. 45, n. 2, 2021, pp. 435-449. DOI: 10.31341/jios.45.2.5
- [8] "Vegetation indices NDVI, EVI, GNDVI, CVI, True color," Vegetation indices ndvi, evi, GNDVI, CVI, True Color. [Online]. Available: <https://www.soft.farm/en/blog/vegetation-indices-ndvi-evi-gndvi-cvi-true-color-140>. [Accessed: 30-Oct-2022].

Алгоритми пошуку шляху для безпілотних сільськогосподарських систем

Валентин Соболев, Андрій Подорожняк

Кафедра комп'ютерної інженерії та програмування
Національний технічний університет “Харківський політехнічний інститут”
Харків, Україна

Анотація—Задачі знаходження шляху зустрічаються в різноманітних сферах людської діяльності – від комп'ютерних ігор і мобільних роботів до логістики перевезень по світу. В подібних задачах дуже рідко найкоротший шлях є оптимальним, оскільки, в залежності від умов задачі, на “ціну” шляху можуть впливати не тільки відстань, а й час, за який ця відстань буде пройдена, витрати, які ми понесемо, та багато іншого. Існує багато алгоритмів для вирішення задач прокладення маршрутів, які мають свої переваги та недоліки. Цікавою є інтерпретація цієї задачі у застосуванні для сучасних та перспективних безпілотних сільськогосподарських систем.

Ключові слова—*Алгоритм пошуку шляху, Оптимізація шляху, Безпілотна сільськогосподарська система.*

I. ВСТУП

Задачі визначення оптимального маршруту зустрічаються як в повсякденному житті людини, так і в процесах великих компаній [1]. Маршрут кур'єра, який розвозить пакунки по адресам, маршрут руху штучного інтелекту в комп'ютерній грі, маршрут перевезень товарів по всій планеті. Всі ці задачі можливо вирішити завдяки алгоритмам прокладення маршрутів [1, 2]. У випадку якщо ви використовуєте мобільний застосунок з картою для визначення свого маршруту по місту, то ефективність цього маршруту – це питання кількох хвилин і сотні зайвих кроків, то у випадку, якщо це велика логістична компанія, перед якою стоїть задача налаштування логістики перевезення по всій планеті, то кожна помилка в розрахунках буде коштувати вам мільйони збитків, яких ви могли уникнути. Особливо актуальна ця задача для сучасних безпілотних сільськогосподарських систем [3]: тракторів, саджалок, протравлювачів, обприскувачів, поливалок тощо.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

На сьогоднішній день існують безліч різноманітних алгоритмів знаходження шляху та їх модифікацій [1, 4], такі як алгоритм Дейкстри, та його модифікація алгоритм A*, алгоритм випадкового блукання, алгоритм Єна та інші. В дослідження та публікаціях, загалом, надаються порівняння роботи схожих алгоритмів та місця використання алгоритмів для рішення задачі знаходження шляху в різних сферах діяльності людини [5].

III. МЕТОДИ

Кожен алгоритм має свої особливості та може бути застосований для вирішення задачі при певних умовах та вхідних даних. Порівняння алгоритмів має сенс лише в випадку, коли вони мають однакові вхідні дані і працюють за схожих умов. Для порівняння алгоритми було поділено на підгрупи, які працюють на схожих вхідних даних. Алгоритми було реалізовано на мові програмування python [6], а програми запущено на одному тому ж самому пристрої, що дає змогу виміряти час, за який алгоритм знайшов рішення, та об'єм пам'яті, який було зарезервовано програмою для виконання роботи. Було підготовлено

декілька вхідних даних різної складності, для того щоб мати змогу оцінити алгоритми при різному обсязі роботи.

IV. РЕЗУЛЬТАТИ

У дослідженні було розглянуто алгоритми пошуку в ширину, алгоритм Дейкстри, жадібний алгоритм пошуку A*, алгоритм випадкового блукання, алгоритм Єна, SSSP-алгоритм, APSP-алгоритм. Отримано дані часу вирішення задач та задіяний об'єм пам'яті для різних наборів вхідних даних на одному тому ж самому обчислювальному пристрої.

V. ОБГОВОРЕННЯ

Отримані результати можна використовувати як підґрунтя для вибору алгоритму для реалізації рішення задачі знаходження шляху для різноманітних безпілотних систем сільськогосподарського призначення.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Для подальшого дослідження слід розглянути інші алгоритми, які сьогодні існують. Для порівняння роботи алгоритмів слід підготувати більшу вибірку даних. Для того, щоб краще розуміти алгоритм та його роботу, слід створити систему, яка б фіксувала стан алгоритму на кожному кроці його роботи.

VII. ВИСНОВКИ

У даному дослідженні розглянуто алгоритми пошуку шляху. Алгоритми було проаналізовано, виявлено їх особливості та місця можливого використання. Для порівняння алгоритми розділено на групи для яких було підготовлено умови для тестування. Алгоритми реалізовано на мові програмування Python та запущено на одному тому ж самому обчислювальному пристрої. Для кожного алгоритму та кожного набору вхідних даних наведено час та об'єм пам'яті, який виділено для виконання алгоритму. Результати, отримані після виконання алгоритму, було порівняно з результатами від інших алгоритмів на тих самих наборах даних.

ЛІТЕРАТУРА

- [1] D. Delling, P. Sanders, D. Schultes, and D. Wagner, "Engineering route planning algorithms" in *Algorithmics of Large and Complex Networks. Lecture Notes in Computer Science*, vol. 5515, Springer: Berlin, 2009, pp. 117-139, doi: 10.1007/978-3-642-02094-0_7.
- [2] M. Needham, and A. E. Hodler, "Graph Algorithms: Practical Examples in Apache Spark and Neo4j," 1st ed., O'Reilly Media, 2019, 257 p.
- [3] V. Sobol, and A. Podorozhniak, "Automation of agricultural machinery" ["Avtomatyzatsiia roboty silskohospodarskoi tekhniki"] in *Problems of informatization Conf.*, Nov. 26-27, 2020. [Online]. Available: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/52853> (last accessed 30.10.2022)
- [4] V. Yaloveha, A. Podorozhniak, and H. Kuchuk, "Convolutional neural network hyperparameter optimization applied to land cover classification," *Radioelectronic and Computer Systems*, no. 1, pp. 115-128, Mar. 2022, doi: 10.32620/reks.2022.1.09.
- [5] Podorozhniak A., Balenko O., and V. Sobol, "Model and algorithms for determining the location and position of agricultural machinery during the movement," *Innov. Tech. and Scient. Solut. for Industr.*, no 2 (16), pp. 32-38, Jun. 2021, doi: 10.30837/ITSSI.2021.16.032.
- [6] M. Lutz, "Learning Python", 5th Edition, O'Reilly Media, 2013, 1643 p.

Мобільний додаток «Система моніторингу асинхронних двигунів» на базі ОС Android

Ольга Чорна, Данило Непорада, Олексій Чорний, Денис Калінін

*Інститут електричної інженерії та інформаційних технологій
Кременчуцького національного університету імені Михайла Остроградського
Кременчук, Україна*

Анотація — В статті розглянуто питання розробки інформаційної технології моніторингу стану асинхронних двигунів за вимірюванням індукції магнітного поля на поверхні статора. Розроблена інформаційна технологія оцінки стану дозволяє проводити попередню діагностику двигуна з видачою рекомендацій про можливість і доцільність його подальшої експлуатації. Для проведення моніторингу розроблено мобільний додаток на базі ОС Android.

Ключові слова — моніторинг асинхронних двигунів, мобільний додаток, операційна система Android.

I. ВСТУП

В останні роки діяльність вітчизняних підприємств ускладнюється нестабільністю роботи в умовах карантину та війни. Крім того, факторами, що негативно впливають на їх роботу є, також, моральний та фізичний знос обладнання. До такого обладнання відносяться перш за все електродвигуни. Вихід з ладу електродвигуна може привести до аварійних ситуацій, що у свою чергу може призвести до значних матеріальних втрат, пов'язаних із простоями технологічних процесів підприємства, ліквідацією наслідків аварії та ремонтом електродвигуна[1].

Щорічно виходить з ладу і ремонтується до 30% парку електричних двигунів (ЕД), що використовуються в промисловості, і до 70%, що використовуються на транспорті. [2]. Більшість електроприводів, що експлуатуються на сучасних підприємствах, складають асинхронні двигуни (АД). Приблизно 30-40 % всіх несправностей в асинхронних двигунах складають дефекти статора, пошкодження елементів ротора – 10 %, пошкодження елементів підшипників – 40 %, інші пошкодження – 12 % [3]. І тому дуже необхідним є виявлення несправності на ранніх стадіях, що може запобігти виникненню серйозних поломок та пошкоджень двигуна.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Сучасні методи діагностики електричних машин розглянуті в роботах: щодо процесів перетворення енергії в електричних двигунах [4], з діагностики і характеристик асинхронних двигунів [5], з моніторингу, оцінки надійності і працездатності АД [6,7], з аналізу магнітного поля [8], з побудови діагностичних систем [9] та інших дослідників. Але широке впровадження енергоощадних технологій в промисловості України стримується через відсутність нових методів моделювання енергетичних об'єктів та відповідного програмно-апаратного забезпечення, що і обумовлює актуальність і значимість досліджень у галузі інтелектуальної інформаційної підтримки моніторингу стану асинхронних двигунів.

Вирішення сучасних завдань діагностики вимагає залучення принципово нових методів та засобів, до них можна віднести діагностування по картині зовнішнього магнітного поля (ЗМП), акустичні і, навіть, органолептичні методи, які передбачають використання органів чуття людини (огляд, прослуховування). При цьому вимір діагностичних параметрів

проводиться безконтактним методом і, що важливо, в робочому режимі. Використання ЗМП як основного, так і додаткового джерела діагностичної інформації дозволяє більш достовірно оцінити технічний стан і прогнозувати відмови електричних машин. Зовнішні ознаки є множиною діагностичних ознак, які відображають взаємозв'язки внутрішніх і зовнішніх параметрів. Такими діагностичними ознаками можуть служити параметри сталого режиму, що змінюються по характерним законам, залежать від місця виникнення і типу первинних несправностей. Кожна конкретна несправність характеризується одним чи декількома зовнішніми ознаками. В одному випадку зовнішня ознака вказує безпосередньо на наявність конкретної несправності, в іншому випадку характеризує несправність тільки побічно. Отже, для діагностики і прогнозування стану електричної машини і привода необхідно знати характер зміни параметрів при різних первинних несправностях і вибрати з них визначальні. Це дозволить проводити попередню діагностику двигуна з визначенням поточного стану двигуна та видачою рекомендацій про можливість і доцільність його подальшої експлуатації. Вказана задача розв'язується шляхом побудови інформаційної технології моніторингу і, на його основі, розпізнавання поточного стану асинхронних двигунів безпосередньо під час виконання технологічного процесу.

Тому метою даного дослідження є – запропонувати принципи побудови мобільного систем діагностики АД у вигляді додатку для гаджетів, що працюють під керівництвом ОС Android. Для досягнення поставленої мети необхідно вирішити наступні завдання:

- запропонувати математичний апарат для розрахунку поточного стану ізоляції обмоток двигуна;
- розробити форми представлення та обробки інформації про двигуни, що діагностуються та результати вимірювання індукції магнітного поля в системі моніторингу на основі мови програмування Kotlin;
- провести експериментальні дослідження шляхом імітаційного моделювання на основі програмного забезпечення Android Studio.

III. МЕТОДИ

В якості методів дослідження в роботі використані аналіз експериментально отриманих даних з вимірювання магнітного поля; методи математичного моделювання стану електромеханічних перетворювачів і чисельного вирішення диференціальних рівнянь; методи імітаційного моделювання роботи розробленого мобільного додатку.

В основу математичного забезпечення системи діагностики покладена залежність відносного значення поточного ресурсу ізоляції обмоток двигуна z від індукції магнітного поля змінного струму, що виміряна на поверхні статора АД [8]:

$$z = 1 - k_b,$$

де k_b – коефіцієнт Буземанової метрики, який розраховується за результатами вимірювань індукції магнітного поля в 16 точках [9]:

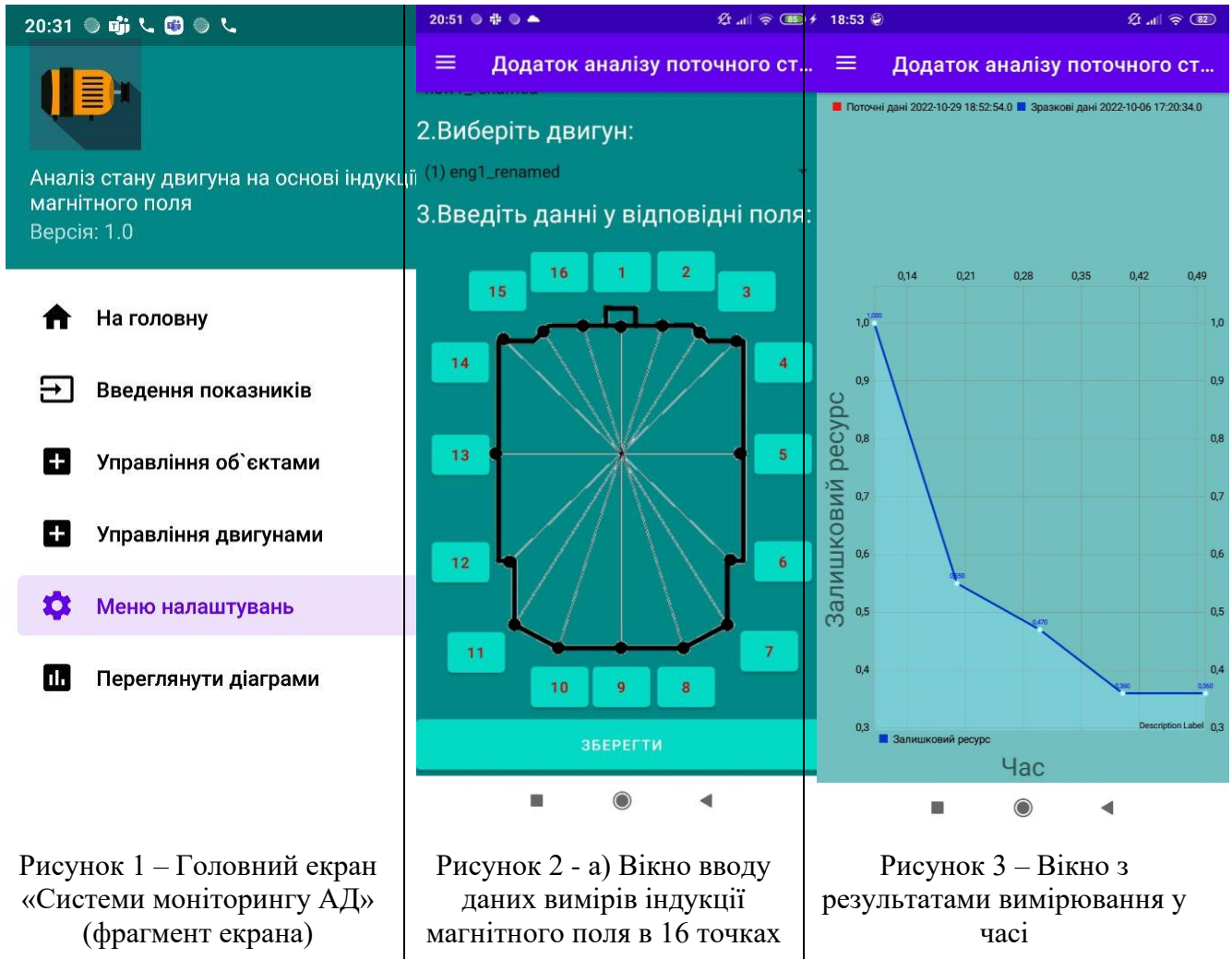
$$k_b = \frac{\left(\sum_{i=1}^N \left| \frac{B_{t,i}}{mB_t} - \frac{B_{m,i}}{mB_m} \right| \right) \left(\sum_{i=1}^N \left| \frac{B_{t,i}}{mB_t} + \frac{B_{m,i}}{mB_m} \right| \right)}{\left(\sum_{i=1}^N \left| \frac{B_{t,i}}{mB_t} \right| \right)^2},$$

де $N=16$ – кількість точок вимірювання, $B_{t,i}, B_{m,i}$ – еталонне і поточне значення виміряної

індукції в i -й точці, мТл, $mB_m = \sqrt{\sum_{i=1}^N B_{m,i}}$, $mB_t = \sqrt{\sum_{i=1}^N B_{t,i}}$.

IV. РЕЗУЛЬТАТИ

Реалізація мобільного додатку виконана на мові Kotlin. Головне вікно, що з'являється на екрані смартфона показане на рис.1, вікна з результатами роботи показані на рис.2 і рис.3.



V. ОБГОВОРЕННЯ

Авторами отримане прикладне вирішення актуальної задачі діагностування стану асинхронних двигунів. Розроблений мобільний додаток дозволяє суттєво спростити процедуру діагностування, регулярне проведення вимірів, наприклад, з початком нової зміни, дозволяє аналізувати оперативну інформацію і завчасно попередити про початок розвитку зміни стану двигуна. До переваг можна віднести те, що даний метод і не потребує зупинки обладнання для проведення діагностики. Запропонований метод і практична його реалізація володіє новизною і наразі не має аналогів у вітчизняній та світовій практиці, адже не потребує складного діагностичного обладнання та витрат часу для проведення діагностики. Розроблена система може використовуватись на будь яких підприємствах з установками обладнаними електричними приводами з асинхронними двигунами.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Система моніторингу дозволяє проводити попередню діагностику двигуна з визначенням поточного стану двигуна та видачою рекомендацій про можливість і доцільність його подальшої експлуатації. Розвиток можливостей і підвищення точності і достовірності

отриманих результатів буде спрямований авторами в напрямку додавання нових діагностичних ознак. В тому числі і органоліптичних.

VII. ВИСНОВКИ

На основі проведеного аналізу систем діагностики електричних машин, реалізації методу «оцінки за станом» для попередньої діагностики стану прийнята діагностична система, що аналізує зовнішні ознаки режиму роботи електричних машин – індукцію магнітного поля змінного струму, що виміряна на поверхні статора АД. Розроблена інформаційна технологія оцінки стану за таким методом дозволяє проводити попередню діагностику двигуна з видачою рекомендацій про можливість і доцільність його подальшої експлуатації. На основі отриманих результатів розроблені основні принципи інтерпретації отриманих результатів для визначення наявності дефектів в двигуні. В основу покладена Буземанова метрика та встановлена лінійна залежність між терміном життя ізоляції обмоток та відносним значенням коефіцієнта Буземанової метрики. Отримані результати дозволяють оцінити термін експлуатації асинхронного двигуна.

Побудований мобільний додаток дозволяє швидко проводити діагностику АД, накопичувати інформацію в базі даних та забезпечує візуалізацію за весь період вимірювань. Впровадження такої системи, на нашу думку, забезпечить значний економічний ефект.

ЛІТЕРАТУРА

- [1] В.М. Кравченко, В.А. Сидоров, *Технічне діагностування механічного обладнання*. Донецьк: Юго-восток, 2010. с. 330.
- [2] О. П. Чорний, Ю. В. Зачепа, В. К. Титюк, О. А. Чорна, *Моніторинг і діагностика електромеханічних об'єктів: навчальний посібник*. Кременчук: ЧП Щербатых А. В., 2019. 122 с.
- [3] А.І. Купін, Д.І. Кузнецов, *Інформаційна технологія для групової діагностики асинхронних електродвигунів на основі спектральних характеристик та інтелектуальної класифікації*. Кривий Ріг: ФОП Чернявський Д. О., 2016. 200 с
- [4] Е. В. Мглинець “Метод діагностування дефектів у розвитку в міжполюсних перемичках при експлуатації синхронних двигунів”, Вісник Національного технічного університету «ХПІ». Серія: Електричні машини та електромеханічне перетворення енергії, №. 3 (1357), С. 56-61., 2020
- [5] С. А. Демішонкова, Я. В. Демішонков, Т. І. Кулік “Функціональна діагностика параметрів трифазного асинхронного двигуна з короткозамкненим ротором”, Вісник Київського національного університету технологій та дизайну. Серія: Технічні науки, №. 4, С. 66-76, 2019
- [6] Д. М. Нестерчук, С. Ф. Курашкін “Діагностування за струмом, як метод захисту електромеханічної системи з асинхронними електродвигунами”, Науковий вісник Таврійського державного агротехнологічного університету імені Дмитра Моторного, Т. 9, №.1, 2019
- [7] В. А. Тытюк, М. Л. Барановская, А. П. Черный, Е. В. Бурдильная, В. В. Кузнецов, К. Н. Богатырев. “Онлайн-идентификация электромагнитных параметров асинхронного двигателя”, Энергетика. Известия высших учебных заведений и энергетических объединений СНГ, 63(5), 423-440, 2020
- [8] M. Zagirnyak, O. Bisikalo, O. Chorna, O. Chorny, "A Model of the Assessment of an Induction Motor Condition and Operation Life, Based on the Measurement of the External Magnetic Field," 2018 IEEE 3rd International Conference on Intelligent Energy and Power Systems (IEPS). 2018. с. 316-321, doi: 10.1109/MEES.2017.8248938.
- [9] О. В. Губаревич, І. В. Мелконова “Імітаційне моделювання асинхронного електродвигуна для підвищення рівня діагностичних систем”, Вісник Східноукраїнського національного університету імені Володимира Даля, №. 1 (271), С. 18-23, 2022
- [10] О. А. Чорна, Д. В. Резнік, Н. В. Зачепа, О. Ю. Кісліцин, Д. О. Кулік, Д. О. Прошин, "Чисельне моделювання розподілу індукції зовнішнього магнітного поля асинхронного двигуна," *Електромеханічні та енергетичні системи, методи моделювання та оптимізації. XVI Міжнародна науково-технічна конференція молодих учених і спеціалістів*, Кременчук, 2018. с. 91-92.

Програмна реалізація сервісу агрегації даних для дослідження комерційних ринкових пропозицій автомобільного ринку України

Тетяна Селівьорстова, Михайло Калюжний, Юрій Каліберда

*Кафедра інформаційних технологій і систем
ННІ «Інститут промислових та бізнес технологій»
Український державний університет науки і технологій
Дніпро, Україна*

Анотація—В роботі описано процедуру проектування та реалізації сервісу агрегації даних для дослідження комерційних ринкових пропозицій автомобільного ринку України. Розглянуті особливості агрегації складних даних. Наведені проектні схеми додатку та розроблений web-інтерфейс сервісу «AutoAggregation».

Ключові слова—Інженерія програмного забезпечення, Комп'ютерні науки, Business Intelligence, Агрегація даних, Автомобільний ринок України

I. ВСТУП

Сучасний бізнес не може бути ефективним без використання відповідних інструментальних засобів аналізу даних [1]. Загалом такі інструменти описуються поняттям Business Intelligence (BI), який спрямований на прийняття ефективних бізнес-рішень шляхом використання відповідних аналітичних інструментів, програмних додатків і технологій для збору, зберігання, аналізу даних і забезпечення ефективного до них доступу (BI-систем).

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Військова агресія росії проти України вплинула на зниження доходів населення, і як наслідок, зменшення кількості продажів автомобілів на первинному ринку, при цьому активність на ринку бувших у використанні автомобілів залишається високою. Це обумовлено зростанням потреби в індивідуальних засобах пересування, які є комфортнішими і безпечнішими в умовах пандемії та військових дій. В умовах конкуренції комерційних ринкових пропозицій важливим аспектом успішності є організація швидкого реагування на ринкові зміни, що в повній мірі стосується автомобільного ринку України. Саме тому задача агрегації комерційних даних автомобільного ринку України має практичне значення, а її реалізація викликає цікавість студентів напрямку «Комп'ютерні науки».

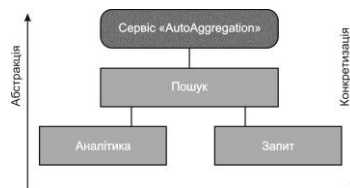
Метою роботи є розробити сервіс агрегації даних для дослідження комерційних ринкових пропозицій на прикладі автомобільного ринку України, який реалізує основні аспекти BI-систем [2]: зберігання, інтеграція, аналіз і представлення даних.

III. МЕТОДИ

В ході розробки сервісу в якості джерел інформації щодо ринкової ситуації з продажу автомобілів було обрані наступні: auto.ria.com, olx.ua, rst.ua (рисуюнок 1).



Рисуюнок 1 – Концептуальна схема розробленого сервісу «AutoAggregation»



Рисуюнок 2 – Блокова структура сервісу «AutoAggregation»

Сервіс реалізує отримання по запиті наступної агрегованої інформації: тип автомобіля; область України, де продається автомобіль; марка автомобіля; модель автомобіля; рік виробництва: від, до; ціна автомобіля: від, до; тип коробки передач.

Крім того, сервіс реалізує видання аналітичної інформації у складі: мінімальна / максимальна / середня ціна в поточному місяці по областях України згідно запити користувача. Реалізовано процедуру генерації повідомлення на електронну пошту у разі появи нової пропозиції згідно пошуку.

IV. РЕЗУЛЬТАТИ

Програмна реалізація концептуальної та блокової структури сервісу «AutoAggregation» з врахуванням основних метрик притаманних автомобільному ринку України дозволили отримати програмний продукт, який відповідає властивостям систем Business Intelligence в частинах зберігання, інтеграція, аналіз і представлення даних (рисунок 3).

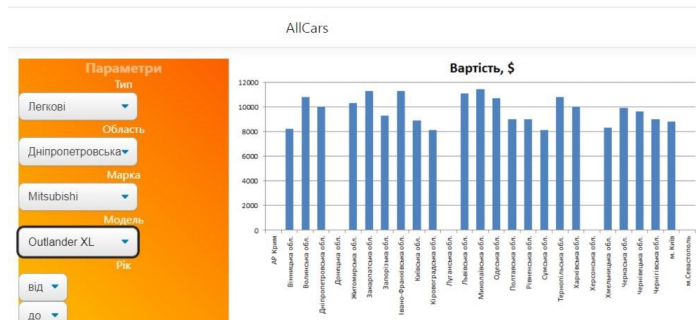


Рисунок 3 – Інтерфейс розробленого сервісу «AutoAggregation»

Розроблений сервіс агрегації даних автомобільного ринку України реалізує процедуру доступу через API до сховищ даних auto.ria.com, olx.ua та rst.ua згідно рисунку 1.

V. ОБГОВОРЕННЯ

Розробка сервісу агрегації даних для дослідження комерційних ринкових пропозицій автомобільного ринку України надала можливість розробити програмний продукт, який відноситься до класу систем Business Intelligence. Крім того, слід зазначити, що реалізовано процедуру доступу до сторонніх сховищ даних, інтеграцію, аналіз і представлення різнорідних даних.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Розроблений сервіс «AutoAggregation», на даний момент, виконує лаконічний обсяг запитів і обробки інформації, що дає перспективи до подальшого розвитку в напрямку розширення метрик Business Intelligence.

VII. ВИСНОВКИ

Розроблений сервісу агрегації даних для дослідження комерційних ринкових пропозицій, реалізує зберігання, інтеграцію, аналіз і представлення даних, що продажів на авторинку України. Крім того, проект «AutoAggregation» цікавий реалізацією агрегування даних через API для доступу до комерційної інформації.

ЛІТЕРАТУРА

- [1] Stanley, Jay (January 2003). "Bigger Monster, Weaker Chains: The Growth of an American Surveillance Society". American Civil Liberties Union.
- [2] H. P. Luhn, "A Business Intelligence System," in IBM Journal of Research and Development, vol. 2, no. 4, pp. 314-319, Oct. 1958, doi: 10.1147/rd.24.0314.

Інтелектуальна технологія синтезу автоматизованої експертної системи керування підприємством

Сергій Альошин, Олена Гайтан

*Кафедра комп'ютерних та інформаційних технологій і систем
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
Полтава, Україна*

Анотація—Робота присвячена інтелектуальній технології синтезу автоматизованої експертної системи керування підприємством, спрямованої на досягнення максимальної ефективності виробничої та фінансово-економічної діяльності при заданих обмеженнях. У роботі розглядається синтез комплексної нейромережевої моделі трьох базових контрольних функцій виробничо-економічної діяльності підприємства, а саме: оцінки актуального стану підприємства, прогнозу динаміки його інформативних показників та розрахунку значень керуючих факторів, адекватних необхідному стану підприємства. Заявлена мета дослідження досягається застосуванням нейромережевих моделей різної архітектури та складності у просторі інформативних факторів-ознак стану підприємства. Навчання моделей реалізується в парадигмі градієнтного методу найшвидшого спуску стосовно функції помилки, що формується. Особливість полягає в тому, що в процесі функціонування нейромережевої моделі реалізується два типи градієнта складної функції. Інструментально синтез моделі реалізується стандартним пакетом технічного аналізу у базисі існуючих градієнтних методів навчання нейронних мереж алгоритмом зворотного розповсюдження помилки. Адекватність нейромережевих моделей встановлюється за продуктивністю та помилками на навчальних та тестових множинах, що дозволяє стверджувати про спроможність прийнятих рішень за результатами моделювання.

Ключові слова—*Експертна система, Нейронна мережа, Підприємство, Керування, Метод найшвидшого спуску.*

I. ВСТУП

Експертна система – комп'ютерна система, здатна частково або повністю замінити фахівця-експерта у вирішенні деякої проблемної ситуації, що імітує здатність експерта-людини приймати рішення [1]. Для управління підприємством експертна система повинна функціонувати в умовах вхідних даних і знань, що змінюються в часі. Саме ця риса експертної системи обумовлює залучення інтелектуальних технологій у базисі штучних нейронних мереж, що мають широкі можливості: в даний час класифікувати актуальний стан підприємства за його об'єктивними показниками, прогнозувати динаміку цих показників у часі і оперативно розраховувати значення керуючих вхідних факторів, відповідних необхідному (цільовому) стану підприємства. Якщо ці три завдання вдається вирішити за допомогою деякої моделі, то успішно вирішується задача синтезу необхідної експертної системи прийняття рішень щодо забезпечення керівництва підприємства необхідними знаннями та даними щодо підтримки необхідного стану підприємства. Це дозволить підприємству мати динамічний та оперативний механізм корекції та утримання заданих показників виробничої, економічної, фінансової, комерційної діяльності на основі безперервного моніторингу показників якості процесу та управління набором залежних факторів для досягнення заданого критерію якості [2]. Існуючі експертні системи вимагають адаптації до особливостей предметної області конкретного підприємства, що не завжди продуктивно [1,2]. Тому, на наш погляд, доцільно здійснити синтез цільової моделі для

конкретного підприємства та врахувати його особливості у процесі реалізації трьох вище названих базових завдань експертної системи – розпізнавання актуального стану, оцінки динаміки показників якості та розрахунку факторів, що управляють, адекватних цільовому стану підприємства.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Використання інтелектуальних технологій для автоматизованого керування підприємством розглянуто в багатьох роботах закордонних та українських дослідників. Так у роботі [3] розглядається використання регресійного та нейромережевого моделювання у моніторингу продукції промислового підприємства. Автори [4] пропонують методичний інструментарій побудови інноваційного середовища підприємства із використанням генетичного алгоритму та нейронних мереж. У даній статті проаналізовано переваги генетичних алгоритмів у пошуку оптимальних рішень порівняно з класичними методами. Автори стверджують, що ключовою проблемою впровадження ефективних промислових інновацій є відсутність сприятливого кліматичного середовища, що стимулює створення інновацій, які забезпечують зростання глобальної конкурентоспроможності, продуктивності праці та якості життя населення. У роботі [5] виділяються особливості ефективного управління підприємством з використанням нейронної мережі та алгоритму зворотного поширення помилки. У роботі [6] пропонується використовувати інструментарій нечітких нейромережевих класифікаторів для моделювання підприємства. Робота [7] описує використання нейронних мережевих архітектур як для оцінки поточного стану підприємства, так і прогнозу його рівня в залежності від фінансового стану підприємства. Автори підкреслюють, що цей продукт є потужним рішенням, спрямованим на отримання прогнозу ефективності управління підприємством.

Враховуючи проведений аналіз останніх досліджень і публікацій, було виділено мету дослідження. Метою дослідження є створення інтелектуальної експертної системи керування підприємством, спрямованої на досягнення максимальної ефективності виробничої та фінансово-економічної діяльності при заданих обмеженнях. Передбачається синтезувати адекватну комплексну нейромережеву модель трьох базових контрольних функцій виробничо-економічної діяльності підприємства, а саме: оцінки актуального (поточного) стану підприємства, прогнозу динаміки його інформативних показників та розрахунку значень керуючих факторів адекватних необхідному стану [8,9,10].

III. МЕТОДИ

Навчання моделей реалізується в парадигмі градієнтного методу найшвидшого спуску стосовно функції помилки, що формується. Особливість полягає в тому, що в процесі функціонування нейромережевої моделі реалізується два типи градієнта складної функції: один формується з часткових похідних функції помилки мережі за її ваговими коефіцієнтами; другий – з часткових похідних функції неузгодженості поточного та цільового стану за множиною вхідних керуючих факторів. Реалізація такого підходу здійснюється із застосуванням синтезу нейромережевої моделі за методом подвійної нейронної мережі (рис. 1).

Масив ознак-факторів мережі $X^n = \{x_1, x_2, \dots, x_n\} \subset X$ разом з алфавітом класів дозволяє реалізувати правило розпізнавання образів:

$$\omega_g \in \Omega_k, \text{ якщо } L(\omega, \{\omega_g\}) = \sup_i L(\omega, \{\omega_i\}), \quad (1)$$

де $X^n = \{x_1, x_2, \dots, x_n\} \subset X$ – ознаки-фактори; $L(\omega, \{\omega_i\})$ – правило віднесення стану мережі ω_g до відповідного класу-стану (банкрут-успішний) у просторі ознак (k, l) за всіх їх можливих поєднань $(\omega_{pk}, \omega_{gl})$.

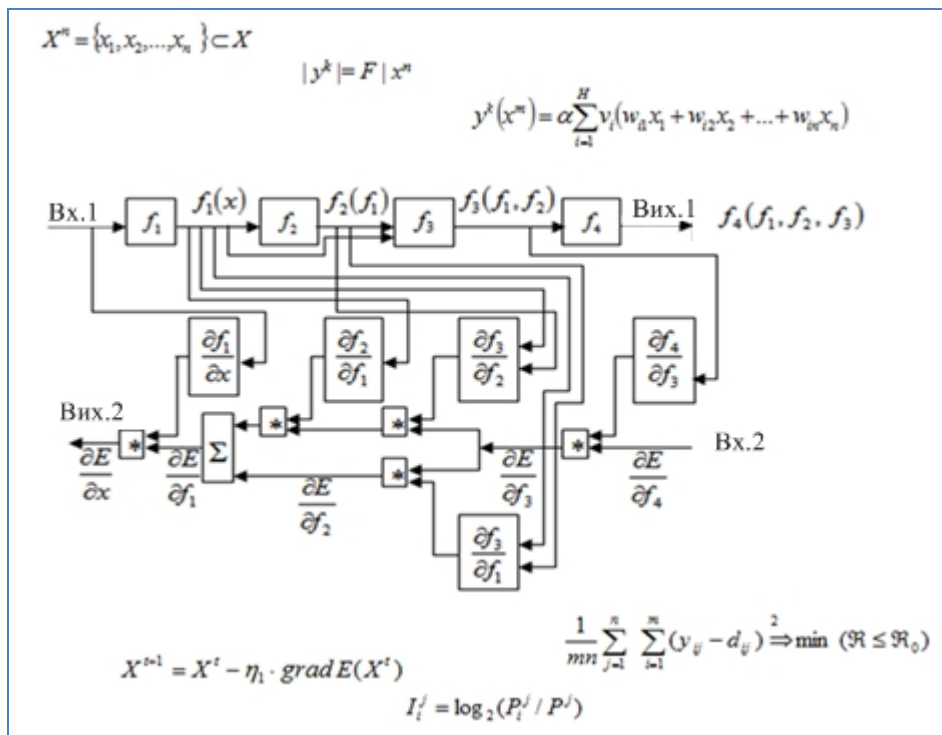


Рисунок1. Схема реалізації аналітичних процедур синтезованої експертної системи

Модифікація синаптичних коефіцієнтів виконується на основі градієнтного метода [2,10]:

$$W^{t+1} = W^t - \eta \cdot \text{grad}E(W^t). \quad (2)$$

та реалізується нейромережовим модулем пакета технічного аналізу [9]:

$$w_{hq}^{(n)}(t) = w_{hq}(t-1) + \Delta w_{hq}^{(n)}(t), \quad w_{hq}(t-1) = w_{hq}(t) + \alpha \frac{\partial E(k)}{\partial w_{hq}(t)}, \quad (3)$$

де w – масив вагових коефіцієнтів мережі; n – номер шару мережі; q – номер виходу нейрона в n -му шарі мережі.

Реалізація цього етапу демонструється виконанням розрахункових процедур, розкритих на схемі (рис. 1) набором опцій у секторі (Вх.1 – Вих.1).

Етап розрахунку вхідних керуючих факторів, адекватних заданій цільовій функції, може бути формально представлений послідовністю процедур:

$$Y_0(t) \rightarrow F_0 : Y_0(t) \rightarrow X_0(t), \quad (4)$$

де $Y_0(t)$ – цільовий вектор показників стану підприємства; $X_0(t)$ – вектор розрахункових значень вхідних факторів; F_0 – функціонал модифікації значень вхідних факторів, адекватних необхідному стану у вигляді:

$$X^{t+1} = X^t - \eta_1 \cdot \text{grad} E(X^t) \quad (5)$$

Реалізація цього етапу наочно демонструється виконанням розрахункових процедур, розкритих на схемі (рис.1) набором опцій у секторі (Вх.2 – Вих.2).

Інструментально синтез моделі реалізується стандартним пакетом технічного аналізу у базисі існуючих градієнтних методів навчання нейронних мереж алгоритмом зворотного розповсюдження помилки [9,10].

Адекватність нейромережових моделей встановлюється за продуктивністю та помилками на навчальних та тестових множинах, що дозволяє стверджувати про спроможність прийнятих рішень за результатами моделювання [9].

Функціональна схема моделі оцінки підприємства представлена на рис. 2:

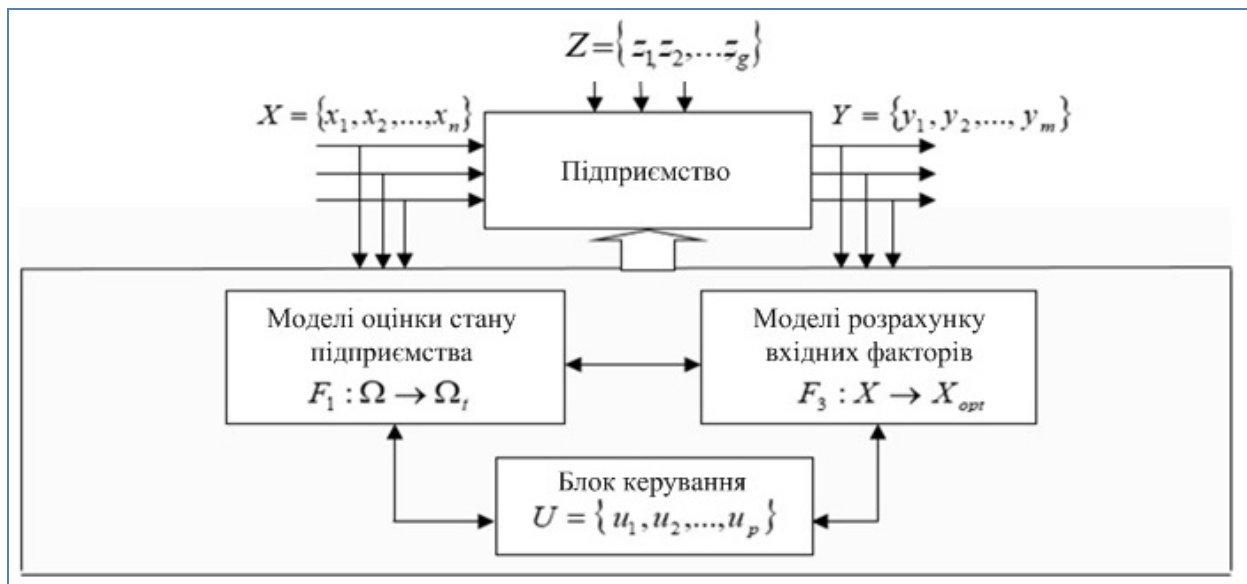


Рис. 2. Моделі базових процесів підприємства

Основними критеріями визначення факту неспроможності підприємства є неплатоспроможність та неоплатність. Як активи виступають оборотні кошти, майно, довгострокові вкладення тощо. Пасивами є залучені кошти та заборгованості за кредитами. Неоплатність виникає у разі, коли пасиви підприємства перевищують вартість його активів.

Для прикладу реалізації експертної системи обрано найбільш поширені ознаки неспроможності підприємства: низька якість наданих послуг або товарів, що продаються; відмова ключових клієнтів від замовлень; зниження обсягу продажу; дисбаланс дебіторської та кредиторської заборгованості; непродуктивна виробничо-комерційна активність організації, що призвела до отримання недостатнього капіталу оборотних коштів; неефективна інвестиційна політика, у результаті якої власного капіталу не вистачає на оборотні кошти; зниження продуктивності використання виробничої потужності, унаслідок чого підприємству довелося збільшити собівартість продукції; спад обсягів виробництва; витрати безпідставно завищені; значно знижено рентабельність продукції; існує висока заборгованість з нездійснених платежів; зросли кредитні суми з урахуванням відсотків; борг, що утворився внаслідок безпідставного збагачення; виникла заборгованість внаслідок заподіяння шкоди майну кредитора; високі неустойки, штрафи, пені за затримку платежів; збитки, що істотно зросли, які підлягають відшкодуванню при невиконанні зобов'язань. Ці фактори узагальнені, стандартизовані та представлені у вигляді навчальної вибірки для ансамблю нейронних мереж різної архітектури та складності [10].

Фрагмент прикладу стандартизованої вибірки наведено у таблиці (таб. 1).

Таблиця 1. Фрагмент прикладу стандартизованої вибірки

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Пер1	Пер2	Пер3	Пер4	Пер5	Пер6	Пер7	Пер8	Пер9	Пер10	Пер11	Пер12	Пер13	Пер14	Пер15	Пер16
-0,5479	0,2919	0,6792	-0,4415	0,6136	1,9925	0,9094	4,0435	0,8837	-0,5479	0,2919	0,6792	-0,4415	0,6136	0,8837	1
-1,0988	-0,6074	0,6520	-0,3159	1,4419	1,7502	0,0522	0,1502	0,5883	-1,0988	-0,6074	0,6520	-0,3159	1,4419	0,5883	2
2,3186	1,4300	0,5840	1,1915	-1,2156	0,4696	0,7104	1,5660	-0,0352	2,3186	1,4300	0,5840	1,1915	-1,2156	-0,0352	1
-0,5173	-0,2702	-0,8574	-0,0088	1,8561	1,0580	-0,4834	-0,4611	-0,2896	-0,5173	-0,2702	-0,8574	-0,0088	1,8561	-0,2896	2
-0,4663	2,1606	-0,0415	0,6053	1,7871	2,6156	-0,3610	0,5203	0,7934	-0,4663	2,1606	-0,0415	0,6053	1,7871	0,7934	1
0,9925	0,8117	-0,2183	1,1915	2,7707	4,1039	1,2767	0,2307	-1,0444	0,9925	0,8117	-0,2183	1,1915	2,7707	-1,0444	2
2,0024	0,0389	-0,3135	0,8844	0,6654	-0,7073	0,3124	2,3704	0,4488	2,0024	0,0389	-0,3135	0,8844	0,6654	0,4488	1
-0,1399	-1,0149	0,5296	-0,1344	-0,1802	-0,4996	-0,5600	-0,6542	0,6621	-0,1399	-1,0149	0,5296	-0,1344	-0,1802	0,6621	2
-0,6499	-0,3685	1,6446	1,4427	-0,1630	-0,7592	0,2971	-0,5255	-0,1911	-0,6499	-0,3685	1,6446	1,4427	-0,1630	-0,1911	1

Процес навчання моделей та вибір найкращих з отриманого ансамблю дозволили зафіксувати архітектуру найбільш продуктивної синтезованої моделі, представленої на рис. 3.

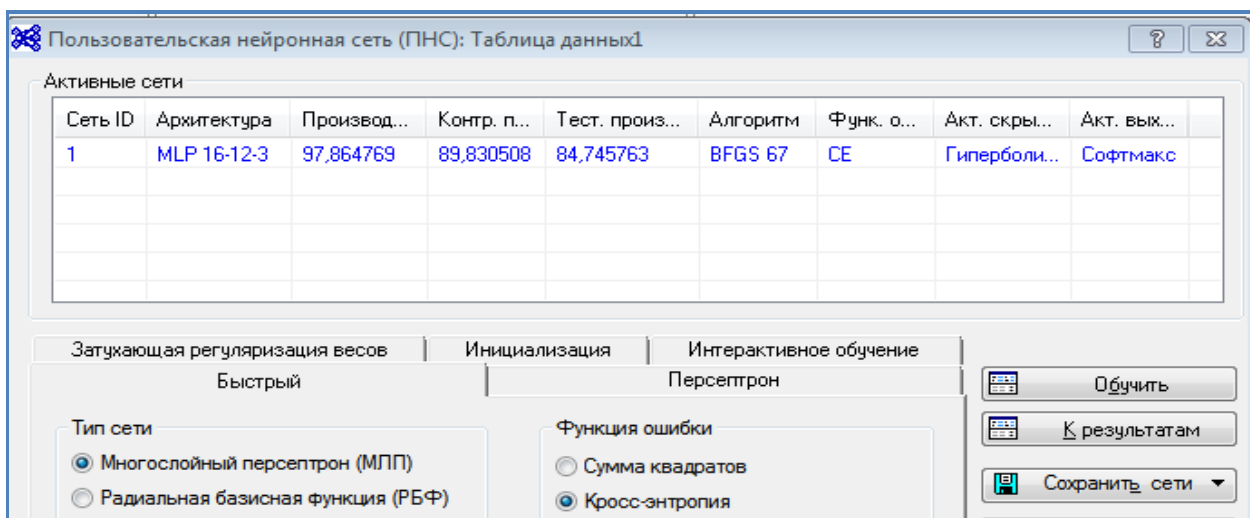


Рисунок 3. Результуюча модель з ансамблю

IV. РЕЗУЛЬТАТИ

Візуалізація процесу збіжності процесу навчання до мінімальних помилок для початкових умов представлена на рис. 4,5:

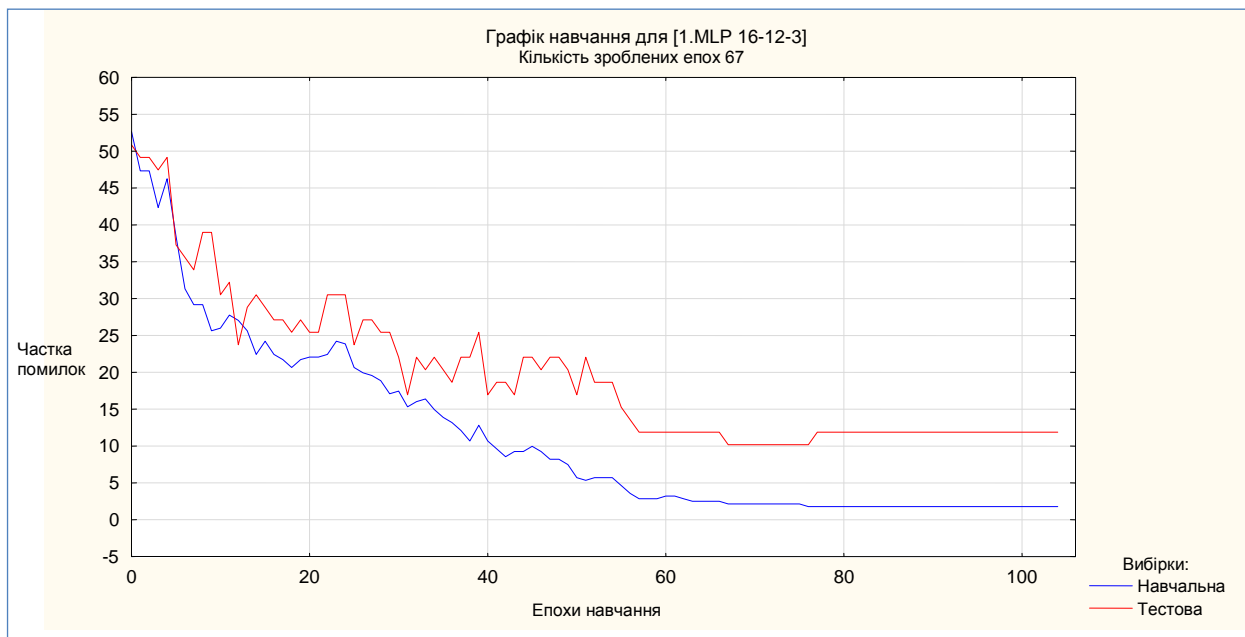


Рисунок 4. Збіжність процесу навчання (варіант 1)

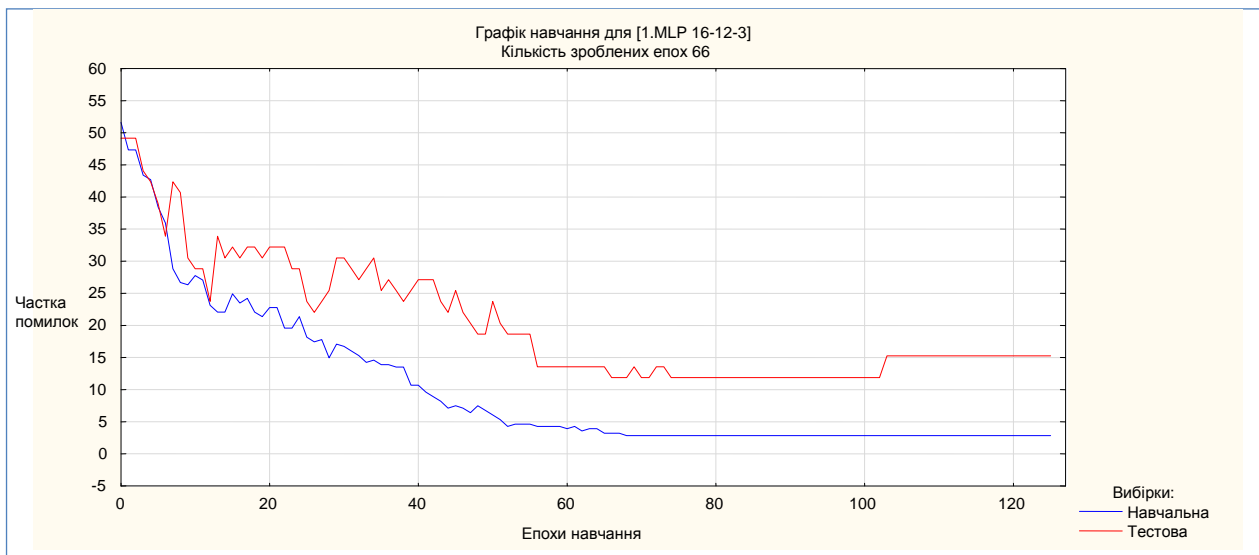


Рисунок 5. Збіжність процесу навчання (варіант 2)

V. ОБГОВОРЕННЯ

Очевидна стійка збіжність процесу навчання до мінімальних помилок у процесі модифікації вагових коефіцієнтів на всій кількості обраних прикладів. Нижча продуктивність на тестовій множині, як видно на рис. 4,5, обумовлена необхідністю подальшого пошуку оптимального співвідношення навчальної та тестової множин, а також вибором часу навчання та продуктивності мережі. Тому вибір кінцевої моделі здійснюється користувачем, виходячи з особливостей предметної області дослідження, вимог до моделей і можливостей нейросередовища в пакеті технічного аналізу даних.

VI. ВИСНОВКИ

Заявлена мета дослідження досягається застосуванням нейромережових моделей різної архітектури та складності у просторі інформативних факторів-ознак стану підприємства для розпізнавання поточного стану, прогнозу динаміки його показників та розрахунку адекватних цільовому стану вхідних факторів. Працездатність запропонованої технології та її ефективність оцінена на репрезентативних даних із наявної бази прикладів.

ЛІТЕРАТУРА

- [1] J.C. Giarratano, G.D. Riley, *Expert Systems: Principles and Programming*, 4th Edition. Course Technology, 2004. 288 pp.
- [2] D. A. Waterman. *A Guide to Expert Systems*. Addison-Wesley, 1986, 419 pp.
- [3] A.R. Kholova, Y.S. Vozhdaeva, etc. "The use of regression and neural network modelling in production monitoring of an industrial enterprise", *Ecology and Industry of Russia*, 25(5), pp. 58-64, 2021 . <https://doi.org/10.18412/1816-0395-2021-5-58-64>
- [4] S. Sviridova, E. Shkarupeta, O. Dorokhova. "The use of neural networks and a genetic algorithm for modeling the innovative environment of enterprises", *E3S Web of Conferences*, 164, 10045, 2020.
- [5] L.Yue, S.Chunyang, "Modeling and optimization of management for enterprise based on artificial neural network", *WIT Transactions on Information and Communication Technologies* 52, pp. 1399-1406, 2014.
- [6] H. Chen, Y.-L. Chen, Applying fuzzy neural network classifier to knowledge management in enterprise modeling. *System Engineering Theory and Practice* 23(11), pp. 1, 2003.
- [7] Д.Настусенко, Я. Ковівчак, "Нейромережові методи та засоби прогнозування показників ефективності управління підприємством", *Технічні вісті*, (49, 50), 64-65, 2020.
- [8] S. Haykin, *Neural Networks and Learning Machines*. Prentice Hall, 2009, 906 pp.
- [9] О.Л. Ляхов, С.П. Альошин, О.О. Бородіна. "Нейромережна модифікація поточного простору ознак до цільової множини класів", *Вісник Донбаської державної машинобудівної академії*, vol. 4 (29), 99-104 pp, 2012.
- [10] Нейромережові засоби штучного інтелекту: навч. Посібник / Р.О.Ткаченко, П.Р.Ткаченко, І.В.Ізонін. - Львів: Вид. Львівської політехніки, 2017. 208 с.

Intelligent system for collecting and analyzing of meteorological data based on robotic platform

Mykhailo Kvasnii¹, Vladyslav Shevchuk¹, Halyna Klym^{1,2}

¹*Department of Specialized Computer Systems
Lviv Polytechnic National University
Lviv, Ukraine*

²*Department of Radioelectronics and Computer systems
Ivan Franko National University of Lviv
Lviv, Ukraine*

Abstract — Systems for collecting meteorological data was investigated. Such system was designed in the form of a quadcopter, which will transmit the collected data in real-time to a server for their systematization and further use through a web application. The structure of the meteorological data collection system has been proposed. The proposed system is based on data collected from drones and isn't tied to a specific place.

Keywords — *Weather Forecast, Drone, Database, Firebase, Data Collection.*

I. INTRODUCTION

The weather was decisive in the survival of people. Nowadays, the study of the weather, and especially its prediction, is an important aspect of the life of society [1]. Modern means of collecting meteorological data mainly include weather stations, meteorological satellites, weather probes, and meteorites. All of them have both advantages and disadvantages. Intelligent robotic platforms in the form of drones are often used to collect and analyze meteorological data. Such meteorodrones are more flexible and cheaper to use, and their main advantage is the ability to collect data without being tied to a specific location [2]. More accurate forecasts require a lot of data from different areas and, preferably, in real-time. Drones, unlike weather stations, can provide this, in addition, they can be used in cases where it is difficult to use the services of meteorologists.

II. RELATED WORK

Many scientists are researching robotic platforms for collecting meteorological data. UAVs can provide high-quality in situ observations of a variety of meteorological variables with high temporal and vertical resolution, and using Raman lidar can provide temperature and humidity profiles with high temporal and vertical resolution in the troposphere [3]. Research has been conducted using an enhanced mobility drone to develop a unified sensor and communication system as a new platform for in situ atmospheric measurements [4]. The main in-situ data base for numerical weather prediction currently relies on radiosonde and airliner observations, with large data gaps above the oceans and in polar regions. These gaps might possibly be patched by measurements with drones, which provide a significant improvement towards environmentally friendly additional data, but so far have not been regarded as a feasible alternative of performing measurements in the upper troposphere [5].

III. METHODOLOGY

Taking into account the peculiarities of predicting weather conditions using modern methods, several main problems can be identified:

- uneven coverage by measuring stations;
- discreteness of observations in time;

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

- processes that are difficult to describe with mathematical equations;
- calculation of available indicators.

Reducing errors when collecting information from the environment will improve the results of processing these values in the future, and obtaining these indicators in real-time will allow measurements to be performed at any moment in time. It is obvious that such a system will not be able to completely solve the problems described above for meteorology in general. However, it will allow you to have actual weather parameters of a certain limited area, which will allow you to adjust actions. Such a system can be useful for farmers, builders, military, or other narrow specialists. It is proposed to develop a system that will allow collecting, processing, and storing weather indicators of a limited area of the territory in real-time from universal drones. The main task of the system is to display data for further processing.

IV. RESULTS AND DISCUSSION

The system should rely on data collected from drones. It has to be configurable and scalable according to the needs of the user. In addition, it shouldn't be tied to a specific place, like a weather station, because it has to collect information from drones. Figure 1 shows a simplified diagram of the entire system. A relational database (DB) should be used to store data. And Realtime Database is used to display data in real-time. Web application connects user with the database. Thus, there is no direct connection between the database and the Realtime Database, which helps to avoid recording unnecessary information. The system should also support other configurations and display capabilities. Final system will consist of several modules that can be quickly modified.

The drone is based on ESP32 board, which has enough computing power, ports supports a large number of data communication protocols, and it's comparably cheap and widely used. The Drone uses BMP680 sensor for data collection, which is an accurate and cheap solution. Drone uses GPS and inertial and ultrasonic sensors, it should be able to navigate in space and make routes according to the given coordinates, and through the communication module, it can directly transmit the collected data to the server.

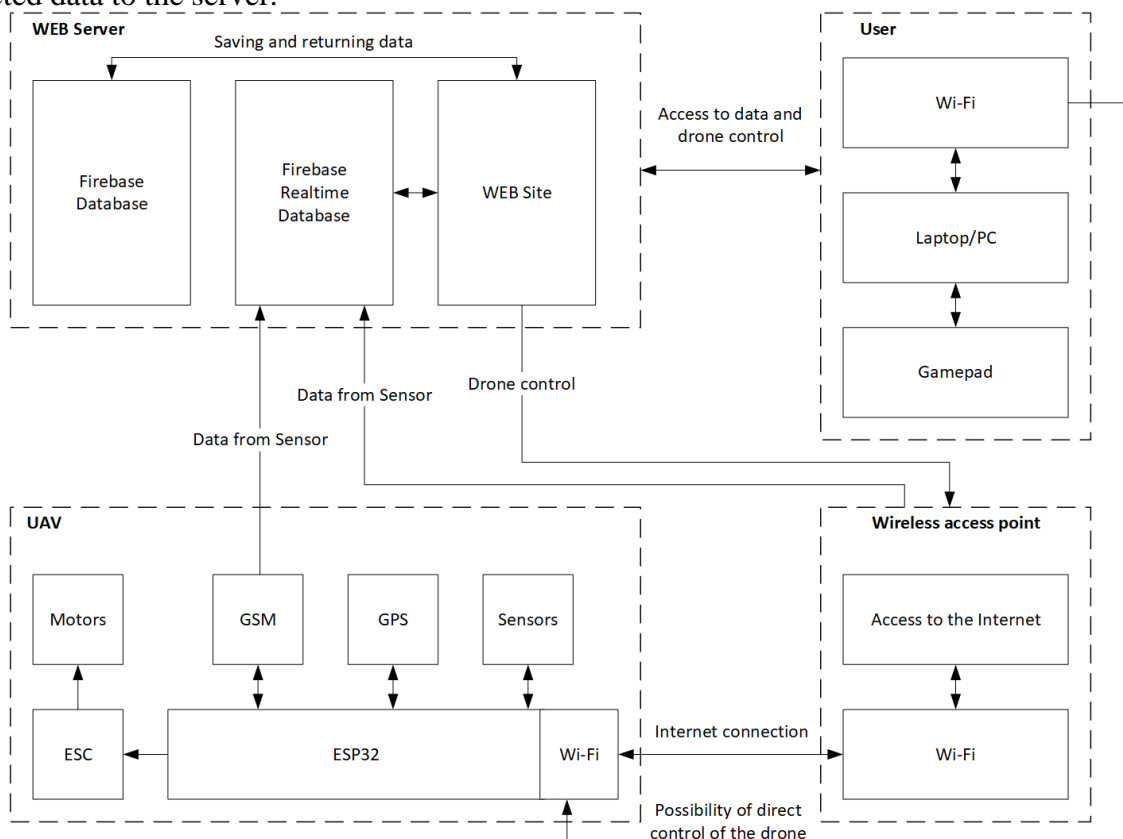


Fig. 1. Outline of the structural diagram of the designed system

The drone has 2 control modes: through web application; The user can be in any part of the world. The drone is connected to the Internet via Wi-Fi; directly via Wi-Fi; Drone is connected to a computer or phone; Data is transmitted to the server via the GSM module.

Through the web application, the user is able to control the drone, set waypoints for the route, and look at the collected data. To control the drone, you can use a game controller, that connects to a computer or phone. The collected information is transferred to the Realtime Database, and after processing, it will end up in Firestore Database, which is a data storage with great productivity, security, and flexible analysis tools.

Through the developed web application, the user can control the drone, set the points for the route, the time and date when the measurements are to be taken and have access to the collected data. By the authorization system, each drone is tied to a user, and the data it collects is labeled accordingly. Added integration with an external API - OpenWeather, which allows you to compare collected data with data collected by weather stations. The drone has a direct control mode for which a game controller is used. The controller must be connected to a computer or phone, and then run the web application. High-precision sensors are used for data collection.

The result of the work is a system capable of measuring and processing weather indicators in a limited area of space in real-time. In this way, we will be able to form a database of historical data that will allow us to conclude climatic changes, as well as partially solve the problems of uneven coverage by weather stations and the discreteness of data in time. The number of simultaneous users and the strict attachment to the data format remains unresolved.

V. FUTURE RESEARCH

A Geiger counter can be added to the drone's sensors to increase the number of measured indicators and to improve the autonomous capabilities of the drone, the ultrasonic sensors can be replaced with lidars. Expanding integration with external APIs, increasing the number of users, and completely delinking from the format of incoming data - all will increase the capabilities of the web application.

VI. CONCLUSION

Designed system for collecting meteorological data, in the form of a quadcopter, which will transmit the collected data in real-time to a server for their systematization and further use through a web application. The structure of the meteorological data collection system has been proposed. The proposed system is based on data collected from drones and isn't tied to a specific place. The system supports authorization and authentication, as well as integration with the OpenWeather API.

VII. Disclosures

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] Y. Shemshuchenko et al., *Legal encyclopedia*, vol. 1, А-Г. Kyiv: Ukrainian encyclopedia named after M.P. Bazhana, 1998.
- [2] E. Ciobanu. "How drones are helping with weather forecasting." Droneblog. <https://www.droneblog.com/how-drones-are-helping-with-weather-forecasting/> (accessed Oct. 31, 2022).
- [3] D. Leuenberger et al., "Improving high-impact numerical weather prediction with lidar and drone observations," *Bulletin of the American Meteorological Society*, vol. 101, no. 7, pp. E1036—E1051, Jul. 2020. Accessed: Oct. 31, 2022. [Online]. Available: <https://doi.org/10.1175/bams-d-19-0119.1>
- [4] H. Madokoro et al., "Development of drone-mounted multiple sensing system with advanced mobility for in situ atmospheric measurement: A case study focusing on PM2.5 local distribution," *Sensors*, vol. 21, no. 14, p. 4881, Jul. 2021. Accessed: Oct. 31, 2022. [Online]. Available: <https://doi.org/10.3390/s21144881>
- [5] K. Bärffuss, H. Schmithüsen, and A. Lampert, "Drone-based meteorological observations up to the tropopause," *Atmospheric Measurement Techniques*, pp. 1–26, Sep. 2022. Accessed: Oct. 31, 2022. [Online]. Available: <https://doi.org/10.5194/amt-2022-236>

Дослідження стаціонарного режиму та перехідного процесу в неідеальних системах

Яна Лебеденко, Юрій Міхлін

Кафедра прикладної математики
Національний технічний університет «Харківський політехнічний інститут»
Харків, Україна

Анотація—Розглядається резонансна поведінка двох систем з трьома степенями свободи, що пов'язані з обмеженим джерелом енергії (або неідеальних систем). Для опису динаміки цих систем використовується метод багатьох масштабів, перехідний процес описано за допомогою апроксимації Паде. Отримані аналітичні розв'язки порівняно з результатами числового моделювання відповідних систем, представлено наближення перехідних процесів до резонансного стаціонарного режиму.

Ключові слова—Системи з обмеженою потужністю, Резонанс, Стаціонарний режим, Перехідний процес, Паде-апроксимації.

I. ВСТУП

Системи з обмеженим джерелом живленням (або неідеальні системи, НІС) характеризуються взаємодією джерела енергії та пружної підсистеми, яка знаходиться під дією джерела. Для неідеальних систем зовнішнє збудження залежить від переміщень пружної підсистеми. Ефект, що проявляється в неідеальних системах – це ефект Зоммерфельда [1], коли в пружній підсистемі виникає режим резонансу з великими амплітудами, і більшість енергії коливань переходить від джерела енергії до цих резонансних коливань. У даній роботі розглянуто резонансну динаміку двох неідеальних систем, а саме системи, яка має маятник як гаситель резонансних коливань, і системи, що описує динаміку порталної рами, яка має вал, що підтримується двома підшипниками, де два сталевих дроти мають істотно нелінійні характеристики. Розглянуті системи показано на Рис.1.а) та 1.б):

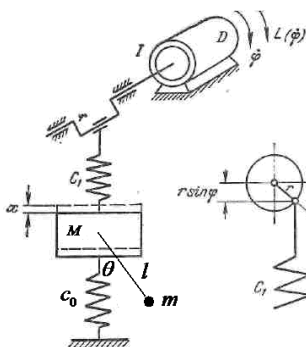


Рисунок 1а). Модель, що містить маятниковий гаситель коливань

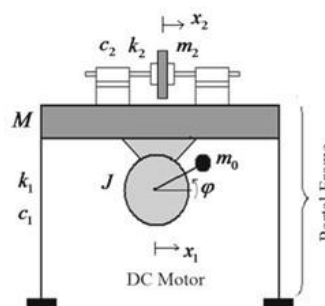


Рисунок 1б). Модель, що містить вал, який підтримується двома підшипниками

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Резонансну динаміку НІС вперше було аналітично описано В.О. Кононенко [2]. Потім дослідження в цьому напрямку були продовжені як Кононенко, так і багатьма іншими авторами. Огляди численних досліджень динаміки НІС можна знайти у статті [3] та у книзі [4]. В публікаціях розглядалися різні види поведінки НІС, включаючи вимушені та

параметричні коливання, автоколивання, перехідні процеси, хаотичні коливання, взаємодію НІС з джерелами енергії різних характеристик тощо. Відомо, що нелінійні гасники коливань дозволяють суттєво зменшити амплітуди резонансних коливань. Зокрема, можливість такого гасіння за рахунок вибору параметрів неідеальної системи обговорюється в статті [5].

III. МЕТОДИ

Рівняння руху для двох даних систем виглядають наступним чином.

Для першої моделі маємо:

$$\begin{cases} (M + m)\ddot{x} + (c_0 + c_1)x = c_1 r \sin \varphi - ml(\ddot{\theta} \cos \theta - \dot{\theta}^2 \sin \theta); \\ I\ddot{\varphi} = a - b\dot{\varphi} + c_1 r(x - r \sin \varphi) \cos \varphi; \\ ml(l\ddot{\theta} + g \sin \theta + \ddot{x} \cos \theta) = 0. \end{cases} \quad (1)$$

Для другої моделі маємо:

$$\begin{aligned} m_1 \ddot{x}_1 + c_1 \dot{x}_1 + c_2(\dot{x}_1 - \dot{x}_2) + k_1 x_1 + \left(\frac{k_2}{l^2}\right)(x_1 - x_2)^3 &= m_0 r(\dot{\varphi}^2 \sin \varphi - \ddot{\varphi} \cos \varphi) \\ I\ddot{\varphi} &= u_1 - u_2 \dot{\varphi} - m_0 r \ddot{x}_1 \cos \varphi \\ m_2 \ddot{x}_2 + c_2(\dot{x}_2 - \dot{x}_1) + \left(\frac{k_2}{l^2}\right)(x_2 - x_1)^3 &= 0 \end{aligned} \quad (2)$$

В першій моделі гасником є математичний маятник, в другій – елемент з малою масою m_2 . В рівняння руху вводиться формальний малий параметр ε , що характеризує малість маси гасника по відношенню до маси основної пружної частини системи, малу дисипацію та малість складових вібрації при зміні в часі кутової швидкості φ по відношенню до основної постійної складової. Введено також для системи (1) нелінійну реакцію пружної підсистеми у вигляді $\varepsilon \tau x^3$ та для системи (2) відповідний член $\varepsilon \lambda x_1^3$. Для опису поведінки систем в області резонансу вводимо також малу частотну «розстройку».

Для опису поведінки систем використовується метод багатьох масштабів [6] та чисельні процедури у пакеті програм Matlab. При побудові стаціонарного розв'язку отримано модуляційні рівняння, що перетворюються на систему нелінійних алгебраїчних рівнянь, яка розв'язується чисельним методом Ньютона. Для побудови аналітичного розв'язку перехідного процесу використовуються дрібно-раціональні апроксимації Паде, що дають змогу описати наближення перехідного процесу до стаціонарного режиму.

IV. РЕЗУЛЬТАТИ

Побудовано стаціонарний режим коливань та перехідний процес в даних системах. Зроблено порівняння стаціонарного розв'язку та перехідного процесу в розглянутих системах. Досліджено вплив зміни параметрів систем на суттєве зменшення амплітуди резонансних пружних коливань, як на малих проміжках часу, так і на достатньо великих проміжках часу., зокрема для системи (1), при зміні параметрів маси маятника m і параметра нелінійності в силі пружності τ . На Рис.2,3 представлено порівняння аналітичного стаціонарного розв'язку з численним моделюванням за допомогою метода Рунге-Кутти 4 порядку. На Рис.2 показано порівняння аналітичного стаціонарного розв'язку з численним моделюванням для системи (1) для змінних x, φ та θ . На Рис.3 показано порівняння аналітичного стаціонарного розв'язку з численним моделювання системи (2) для змінних u, φ та v .

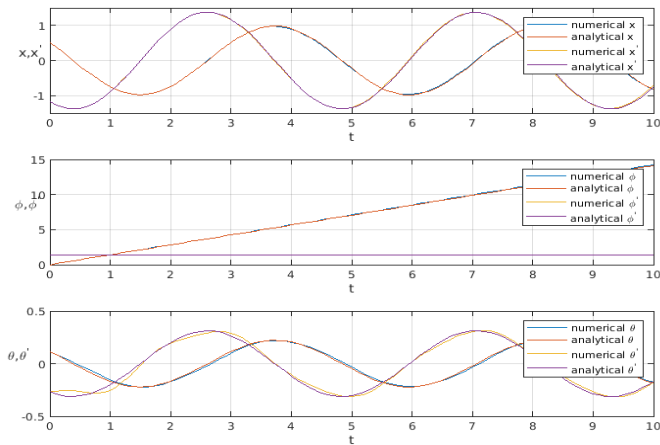


Рисунок 2. Порівняння аналітичного стаціонарного розв'язку з численним моделюванням для системи (1)

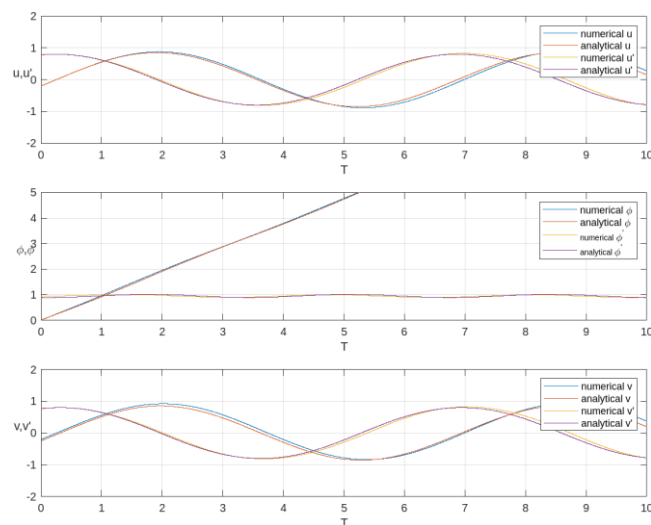


Рисунок 3. Порівняння аналітичного стаціонарного розв'язку з численним моделюванням для системи (2)

V. ВИСНОВКИ

За допомогою методу багатьох масштабів з достатньо великою точністю отримано стаціонарний резонансний режим коливань систем, що розглядається. Використання апроксимацій Паде дозволяє ефективно описати перехідний процес, що наближається до стаціонарного режиму зі збільшенням часу. За допомогою чисельного моделювання показано, що шляхом вибору параметрів системи можна отримати істотне зменшення амплітуд резонансних коливань.

ЛІТЕРАТУРА

- [1] A. Sommerfeld, "Beiträge zum dynamischen ausbau der festigkeitslehre," *Physikal Zeitschr.*, 1902. vol. 3, pp. 266–286, 1902
- [2] V.O. Kononenko, *Vibrating Systems with Limited Power Supply*. London: Illife Books, 1969.
- [3] Balthazar J.M., et al., "An overview on the appearance of the Sommerfeld effect and saturation phenomenon in non-ideal vibrating systems (NIS) in macro and mems scales," *Nonlinear Dynamics*, vol. 93, no 1, pp. 19–40, 2018
- [4] L. Cveticanin, M. Zukovic, J.M. Balthazar, *Dynamics of Mechanical Systems with Non-Ideal Excitation*. Cham: Springer, 2018
- [5] Y.O. Lebedenko, Y.V. Mikhlin, and M.A. Pinsky, "Resonance Dynamics of the Non-ideal System Having the Pendulum as Absorber of Elastic Vibrations," in *Nonlinear Vibrations Excited by Limited Power Sources. Mechanisms and Machine Science*, vol. 116 J.M. Balthazar, Ed. Cham: Springer, 2022, pp. 139-149
- [6] Nayfeh A.H., Mook D.T. *Nonlinear Oscillations*. New York: Wiley, 1979

Прогнозування часу доставки даних в ІТС залізничного транспорту з використанням нейронечіткої технології

Вікторія Пахомова

*Кафедра електронних обчислювальних машин
Український державний університет науки і технологій
Дніпро, Україна*

Анотація—Запропонована модифікація структури інтегрованої системи маршрутизації в інформаційно-телекомунікаційній системі (ІТС) залізничного транспорту, основу якої складають нейронечіткі мережі (ННМ), додаванням ННМ0 (для прогнозування затримки на маршрутизаторі), на вхід якої подаються затримки за попередні п'ять часових проміжків, та ускладненням ННМ1 (для прогнозування часу доставки даних за маршрутом), на вхід якої подаються наступні параметри: кількість проміжних маршрутизаторів, що складають маршрут (10 терм); інтенсивність трафіка (3 терми); сумарна затримка на маршрутизаторах (3 терми); довжина пакета (3 терми). Відповідні вибірки сформовані на основі створеної в Ornet Modeler імітаційної моделі мережі розглянутого фрагменту ІТС залізничного транспорту, для якого здійснений прогноз часу доставки даних за двома маршрутами його проходження (найдовшого та найкоротшого) на основі ННМ0 та ННМ1 конфігурацій 5-10-32-32-1 та 4-19-30-30-1 відповідно, що створені в MatLAB. Проведено дослідження похибки ННМ1 при різних функціях належності та за різними методами оптимізації навчання. Виявлено, що найменше значення середньої похибки надає ННМ1 при Гаусовій функції належності за гібридним методом.

Ключові слова—ІТС, Прогнозування, Маршрутизатор, Затримка, Імітаційна модель, Нейронечітка мережа, Конфігурація, Функція належності, Метод оптимізації, Похибка.

I. ВСТУП

Сучасні алгоритми вибору найкоротшого маршруту, наприклад, алгоритми Беллмана–Форда та Дейкстри, які в даний час продовжують широко використовуватися в протоколах маршрутизації (RIP, OSPF та ін.), не завжди призводять до ефективного результату. Тому виникає необхідність дослідження можливості організації маршрутизації в комп'ютерних мережах, що лежать в основі інформаційно-телекомунікаційної системи (ІТС) залізничного транспорту, за допомогою методів штучного інтелекту.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

На сучасному етапі пошук розв'язку задачі маршрутизації в комп'ютерних мережах на основі використання нейронних мереж проводять різні вчені [1-12]. У свій час першу таку спробу здійснив Хопфілд для розв'язання задачі комівояжера [1]. Павленко М. А. виконав аналіз можливостей деяких нейронних мереж: багаточарового персептрон, мережі Хопфілда, мережі RBF для організації маршрутизації із п'яти маршрутизаторів [2]. У свою чергу, нейронечіткі мережі (ННМ) покликані поєднати в собі переваги нейронних мереж і систем нечіткого виведення [3-5], зокрема адаптивна нейронечітка система (Adaptive-Network-Based Fuzzy Inference System, ANFIS), що реалізована в додатку Fuzzy Logic Toolbox системи MatLAB. Раніше автором в роботах [5-8] були представлені результати досліджень використання інтелектуальних засобів у мережі ІТС залізничного транспорту. Визначення оптимального маршруту в комп'ютерній мережі можна провести на основі прогнозування часу доставки даних за маршрутом, зокрема в [8] автором запропонована структура

інтегрованої системи маршрутизації для Придніпровської залізниці. Але відомо, що донедавна робота залізничного транспорту України являла собою взаємодію шести залізниць: Східної, Львівської, Придніпровської, Харківської, Одеської, Донецької, на кожній з яких впроваджена відповідна ІТС [9], що потребує проведення додаткових досліджень.

Метою даної роботи є модифікування запропонованої структури інтегрованої системи маршрутизації в ІТС залізничного транспорту з використанням нейронечіткої технології.

III. МЕТОДИ

Для вирішення задачі обрання оптимального маршруту застосований метод прогнозування часу доставки пакету за маршрутом (від вузла джерела до вузла призначення). Джерелами затримки передачі пакету у комп'ютерній мережі є: затримка у лініях передачі; затримка, що вноситься самим маршрутизатором; час простою пакету у черзі на маршрутизаторах. Основну затримку вносять черги на маршрутизаторах у найбільш завантажених місцях комп'ютерної мережі. Для організації маршрутизації в ІТС залізничного транспорту пропонується модифікація структури [8], що відображена на рисунку 1.

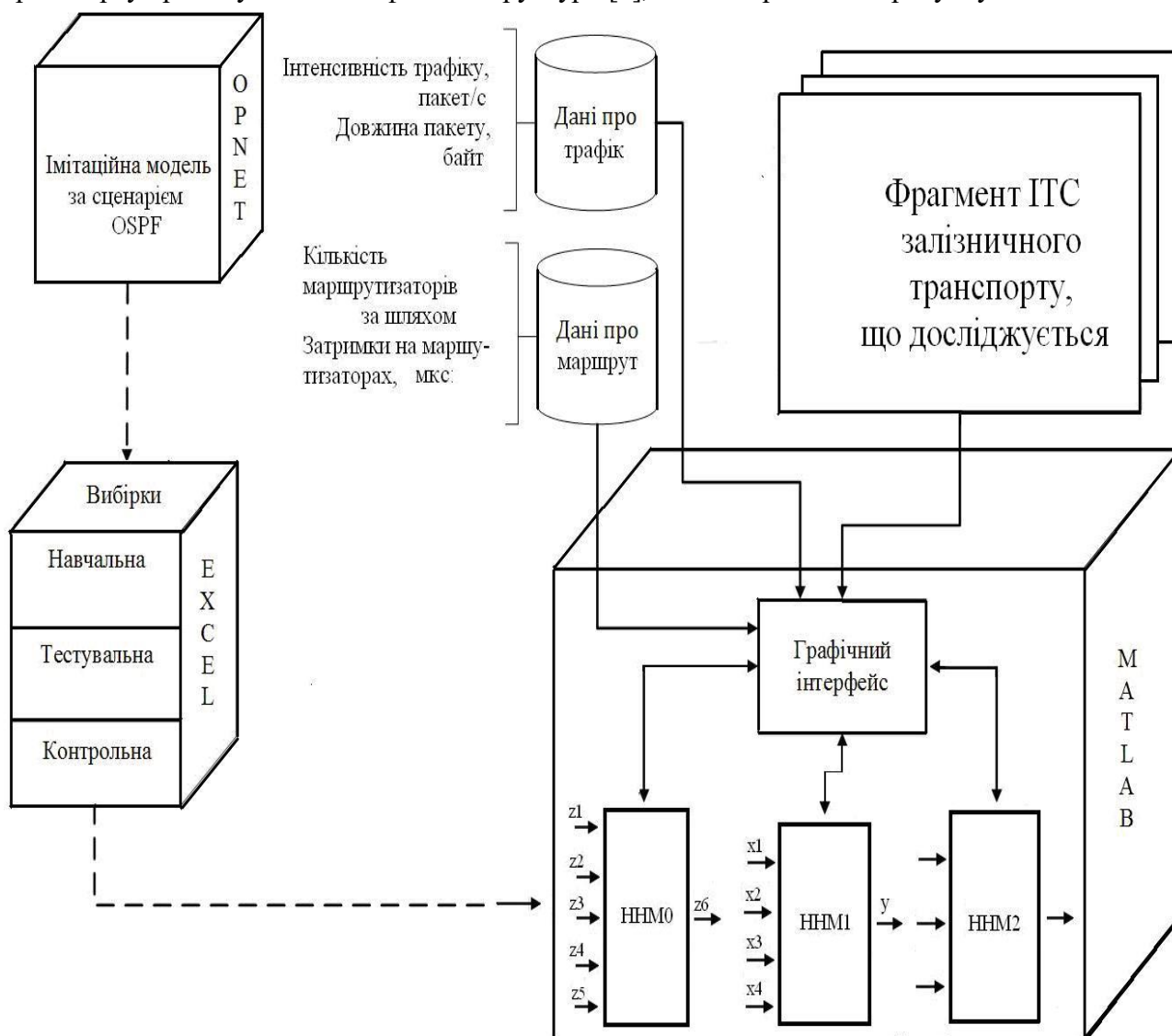


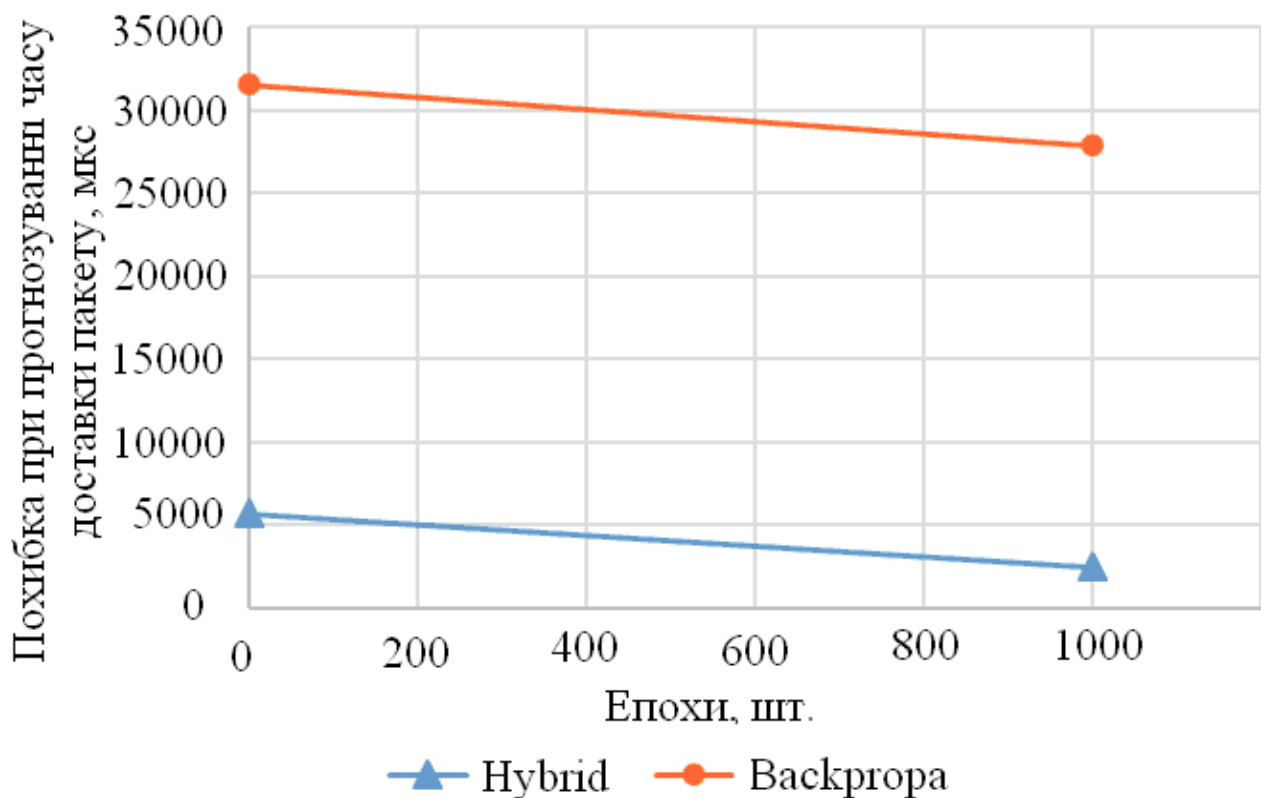
Рисунок 1 – Модифікована структура системи маршрутизації в ІТС залізничного транспорту

У [8] автором була запропонована структура інтегрованої системи маршрутизації в ІТС Придніпровської залізниці, основу якої складають наступні ННМ: ННМ1 конфігурації

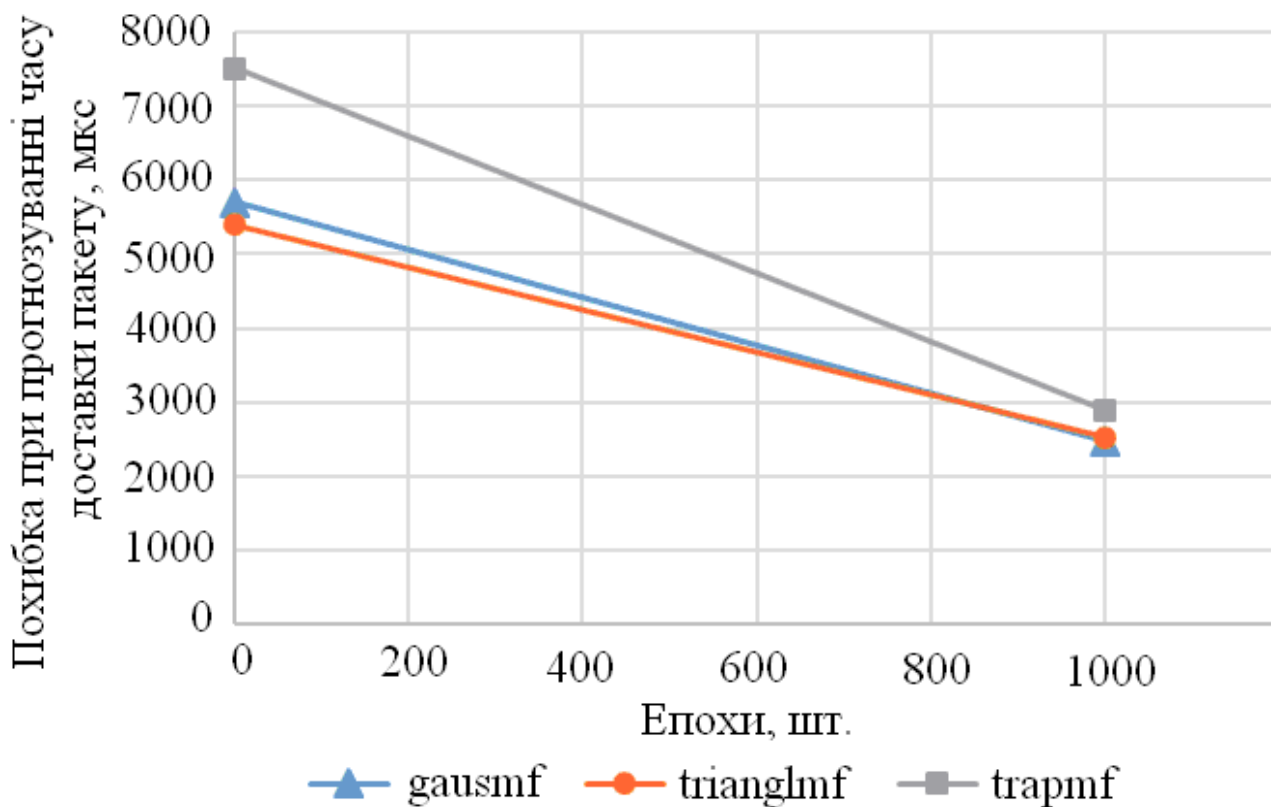
3-12-60-60-1 (де 3 – кількість нейронів шару input, 12 кількість нейронів шару inputmf, 60 – кількість нейронів шару rule, 60 – кількість нейронів шару outputmf, 1– кількість нейронів шару output) для прогнозування часу доставки даних; ННМ2 конфігурації 3-13-75-75-1 (де 3, 13, 75, 75 та 1 – це кількість нейронів першого, другого, третього, четвертого та п'ятого шарів відповідно) для визначення оптимального маршруту за однаковою кількістю маршрутизаторів. Відмінність від [8] стосується додавання ННМ0 (для прогнозування затримки на маршрутизаторі) та ускладнення ННМ1, що створені в MatLAB за алгоритмом Сугено. Якщо відомі значення затримок на маршрутизаторі $\{z(t_1), z(t_2), \dots, z(t_5)\}$ в послідовні часові проміжки t_1, t_2, \dots, t_5 , тоді задача прогнозування затримки на маршрутизаторі полягає в прогнозі значення $z(t_6)$ в майбутній часовий проміжок t_6 . Прогнозування затримки на маршрутизаторі здійснювалося на основі ННМ0 конфігурації 5-10-32-32-1, для лінгвістичної оцінки кожна вхідна змінна якої мала два терми (максимальне та мінімальне значення). Для прогнозування часу доставки пакету в комп'ютерній мережі використовувалася ННМ1 конфігурації 4-19-30-30-1, за вхідні параметри якої використані наступні змінні: x_1 – кількість проміжних маршрутизаторів, що складають маршрут проходження пакета (K_1, K_2, \dots, K_{10}); x_2 – інтенсивність трафіка ($A_{\min}, A_{\text{avg}}, A_{\max}$); x_3 – сумарна затримка на маршрутизаторах ($Z_{\min}, Z_{\text{avg}}, Z_{\max}$); x_4 – довжина пакета ($L_{\min}, L_{\text{avg}}, L_{\max}$). Дані для вибірок знімалися на створеній імітаційній моделі мережі фрагменту ІТС залізничного транспорту (відмінного від мережі Придніпровської залізниці), що побудована в системі Ornet Modeler. За допомогою пакету Microsoft Excel сформовані вибірки: навчальна (720 прикладів), тестувальна (360 прикладів) та контрольна (270 прикладів).

IV. РЕЗУЛЬТАТИ

Дослідження похибки ННМ1 за епохами проводились для різних методів оптимізації навчання: Backpropa (метод зворотного поширення помилки); Hybrid (гібридний метод), і різних функцій належності: Гаусової; трикутної; трапецієподібної (рисунк 2).



(a)



(б)

Рисунок 2 – Дослідження похибки ННМ2 за епохами: (а) – при різних методах оптимізації (функція gaussmf); (б) – при різних функціях належності (гібридний метод)

Із рисунку 2(а) видно, що для ННМ2 початкова похибка становила 5700 мкс та за 10000 епох впала до 2468 мкс при використанні Гаусової функції за гібридним методом, тобто зменшилась у 2,3 рази.

Із рисунку 2(б) видно, що найбільше змінення похибки сталося при застосуванні трапецієподібної функції належності у 2,5 рази, проте початкова похибка була більшою, ніж при використанні інших функцій належності, і складала 7516 мкс.

V. ОБГОВОРЕННЯ

Проведено зрівняння отриманих результатів (часу доставки даних) на імітаційній моделі мережі для фрагменту ІТС залізничного транспорту (відмінного від Придніпровської залізниці) та з використанням ННМ1 конфігурації 4-19-30-30-1 для двох маршрутів: найдовшого (похибка прогнозування часу доставки даних на ННМ1 склала 6 %) та найкоротшого (похибка прогнозування часу доставки даних на ННМ1 склала 3 %).

З одного боку, конфігурація 4-19-30-30-1 в зрівнянні з конфігурацією 3-12-60-60-1 надає можливість врахувати затримку на маршрутизаторах у якості додаткового параметра (що в свою чергу призведе до отримання більш точного результату), з іншого боку, зменшена в два рази кількість правил для її навчання (що в свою чергу має значний вплив на час роботи).

V. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

На наступному етапі необхідно провести додаткові дослідження структури та оптимальних параметрів ННМ0 для прогнозування затримки на маршрутизаторі, а також апробацію повного комплексу ННМ0...ННМ2 з використанням реальних значень затримок.

VI. ВИСНОВКИ

Одним із підходів щодо організації маршрутизації в інформаційно-телекомунікаційній системі залізничного транспорту являється використання комплексу нейронечітких мереж (ННМ): ННМ0 (прогнозує затримку на маршрутизаторі); ННМ1 (прогнозує час доставки даних за маршрутом); ННМ2 (визначає оптимальний маршрут за рахунок однакової кількості маршрутизаторів).

За допомогою додатку Fuzzy системи MatLAB створені ННМ0 та ННМ1 конфігурацій 5-10-32-32-1 та 4-19-30-30-1 відповідно. Навчання ННМ1 проводилося протягом 10000 епох з використанням вибірки із 720 прикладів, затрачений час склав 45 хв.; при навчанні вдалось знизити рівень похибки ННМ2 (від 5700 мкс до 2468 мкс).

Виявлено, що найменше значення середньої похибки ННМ1 досягається при Гаусовій функції належності за гібридним методом оптимізації.

VII. ПОДЯКА

Подяка здобувачам ступеня «магістр» спеціальностей «Комп'ютерна інженерія» і «Кібербезпека» за проведення експериментів на створеній в системі Ornet Modeler моделі комп'ютерної мережі для розглядаемого фрагменту інформаційно-телекомунікаційної системи залізничного транспорту та формуванню відповідних вибірок для прогнозування часу доставки даних на основі ННМ1.

ЛІТЕРАТУРА

- [1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", *Proceedings of National Academy of Sciences*, 1982, vol. 79, iss. 8, pp. 2554–2558, doi: 10.1073/pnas.79.8.2554
- [2] М. А. Pavlenko, "Analysis opportunities of artificial neural networks for solving single-path routing in telecommunication network", *Problemy telekomunikatsii*, № 2(4), retrieved from <http://pt.journal.kh.ua/index/0-139>
- [3] K. Sasikala, and V. Rajamani, "A neuro fuzzy based conditional shortest path routing protocol for wireless mesh network, *International Journal of Enhanced Research in Management & Computer Applications*", 2013, vol. 2, iss. 5, pp. 1–10
- [4] M. V. Kumar, and Dr. T. Lalitha, "Soft Computing: Fuzzy Logic Approach in Wireless Sensors Networks", *Circuits and Systems*, 2016, vol. 07, iss. 08, pp. 1242–1249, doi: <https://doi.org/10.4236/cs.2016.78108>
- [5] V. M. Pakhomova, "Network traffic forecasting in information-telecommunication system of Prydniprovsk railways based on neuro-fuzzy network", *Science and transport progress*, 2016, № 6 (66), pp. 105–114, doi: 10.15802/stp2016/90485
- [6] V. M. Pakhomova, and I. D. Tsykalo, "Optimal route definition in the network based on the multilayer neural model", *Science and transport progress*, 2018, № 6 (78), pp. 126–142, doi: 10.15802/stp2018/154443
- [7] V. M. Pakhomova, T. I. Skaballanovich, and V. S. Bondareva, "Intelligent routing in the network of information and telecommunication system of railway transport", *Science and transport progress*, 2019, № 2 (80), pp. 77–90, doi: 10.15802/stp2019/166092
- [8] V. M. Pakhomova, and Y. S. Mandybura, "Optimal route definition in the railway information network using neural-fuzzy models", *Science and transport progress*, 2019, № 5 (83), pp. 81–98, doi: 10.15802/stp2019/184385
- [9] V. M. Pakhomova, "Doslidzhennia informatsiino-telekomunikatsiinoi systemy zaliznychnoho transportu z vykorystanniam shtuchnoho intelektu: monohrafiia", Dnipro: Standart-Servis, 2018, 220 p, isbn: 978-617-7382-14-9
- [10] K. V. Kolesnikov, A. R. Karapetian, V. Y. and Bahan, "Analiz rezultativ doslidzhennia realizatsii zadachi marshrutyzatsii na osnovi neuronnykh merezh ta henetychnykh alhorytmiv", *Visnyk Cherkaskogo derzhavnogo tehnologichnogo universitetu, Seria: Tehnichni nauky*, 2016, № 1, pp. 28-34
- [11] A. M. Bryndas, P. I. Rozhak, N. O. Semenishin, and R. R. Kurka, "Implementing of the Problem of Choosing the Optimal Flight Rout by a Hopfield Neural Network", *The Scientific Bulletin of UNFU*, 2016, 26.1, pp. 357-363
- [12] W. H. Schuler, C. J. A. Bastos-Filho, and A. L. I. Oliveira, "A novel hybrid training method for hopfield neural networks applied to routing in communications networks", *International Journal of Hybrid Intelligent Systems*, 2009, vol. 6, iss. 1, pp. 27–39, doi: 10.3233/his-2009-0074

Using Microsoft Family Safety for parents and children

Vasyl Franchuk¹, Nataliia Franchuk^{1,2}

¹*National Pedagogical Dragomanov University*

²*Institute for Digitalisation of Education of the National Academy of Educational
Sciences of Ukraine
Kyiv, Ukraine*

Abstract—The article discusses ways to ensure the information security of children and to organize a secure personal information space both at school and in the family. Describes the software for parental control Microsoft Family Safety, and its basic principles of setting up and managing this software for parents and children. The benefits of using parental controls on mobile devices are presented. The main possibilities for using the software for parents and children are described. Outlines the basic steps for creating a child's Google account with parents and managing it from any device on the Internet. Attention is drawn to the fact that children are increasingly looking for entertainment using mobile devices, sharing new products with their peers, and spending time together playing various online games, and parents cannot control this process. At the same time, teachers conduct lessons using mobile devices to draw attention to the subject being studied. In the process of using computer technology, students can form a special type of thinking, the so-called "clip consciousness", which leads to fragmentary thinking and the irreversibility of changes in consciousness. Therefore, the problem of excessive use of mobile devices by children requires constant analysis and coverage. The development, creation, and use of software for monitoring the presence of children on gadgets is one of the urgent needs of parents and teachers. In turn, teachers, in addition to teaching students, need to draw their parents' attention to the use of gadgets outside of school. To do this, parents are encouraged to use a variety of software, including Microsoft Family Safety, which provides parents with tools to monitor their child's mobile and Internet behavior, and helps parents choose child-friendly content and explain how to use their device and Internet resources.

Keywords—*Information Security, Parental Control, Microsoft's Account, Microsoft Family Safety.*

I. INTRODUCTION

Problem statement and substantiation of urgency. In the process of technology development, it becomes necessary to use a certain environment for family members. Today, each of us, and especially children, have gained easy access to the Internet through the use of modern gadgets, which can quickly become addictive to the uncontrolled use of certain applications or due to exposure to unwanted content. Children are increasingly looking for entertainment using mobile devices, sharing new products with their peers, and spending time together playing a variety of online games.

Teachers began to conduct lessons using mobile devices to draw attention to the subject. It should be understood, that the informatization of education is not a panacea, there are risks of loss of cultural and creative thinking, live communication, and the illusion of accessibility of cognitive actions. In the process of using computer technology, students form a special type of thinking, the so-called "clip consciousness", which leads to fragmentary thinking and the irreversibility of changes in consciousness. This causes a bad mood, irritability, often general malaise, or fatigue. As a result, appetite is reduced, sleep is disturbed, and efficiency is reduced (reduced working capacity) [7].

However, with a harmonious combination, the educational process improves, because there is always a lively discussion between students. Researchers have shown that children of all ages memorize learning material best while playing. Therefore, motivated use of mobile devices during learning increases student activity [1].

Formulation of goals and objectives of the article. The problem of the overuse of mobile and stationary devices requires constant analysis and coverage. It is clear, that the development, creation,

and use of software to control the time spent by children on the devices is one of the urgent needs of parents and teachers to prevent the "collage of modern thinking" of children. Formation of awareness about the effective use of existing information technologies in order to self-organize the educational process for the use of various educational tasks through development programs, educational and game complexes, etc.

Purpose of writing the article: to acquaint teachers, parents, and children with the main possibilities of using the software for parental control Microsoft Family Safety.

II. RELATED WORK

Information technologies are increasingly integrated not only with production, but also with education, family life and upbringing. Expanding the age range of users and teaching "from the cradle" to work at a computer, you need to consider pedagogical and educational problems faced by each family. The acquaintance of the younger generation with the principles of working on a computer is not limited to the skills of working on it, but also captures the sphere of family upbringing and communication, as it is part of every family's life from the first days of a child's life. Parents' efforts should be aimed at creating a safe, socially useful, and critical attitude of the child to the computer and the Internet [8].

However, the issue of information and computer security in afterschool and family education is given insufficient attention. Particular attention should be paid to games using mobile and stationary devices. Uncontrolled use of such games affects the child's consciousness, causing cruelty, anxiety, emotional imbalance, and irritability.

However, due to a number of these dangers, a total ban on children using a gadget, tablets, computers, or the Internet is unjustified. These are just tools, and as a result, of motivated use, these devices can be useful assistants in the education and upbringing of the child. Therefore, the task is to teach children to use computers properly, including mobile devices and the Internet. One way to ensure such information security for children is to organize a secure personal information space both at school and in the family. It is possible to organize such a safe information space by implementing measures and information security for children. Among which:

- *legal means of information security of children on a legislative basis*, implemented for a unified state policy in the field of protection of children from malicious information materials, the revision of which harms their health and psyche;
- *ethical and moral measures*, including compliance with the norms and rules of behavior of children in the process of information activities, as well as network culture with the use of information technology;
- *software and technical measures* that involve the use of various hardware and software to prevent material and moral harm to children (parental control programs, technical means of data protection);
- *organizational measures to control* the use of online communities and services, which may preclude harm to the child's information environment;
- *educational activities*, the formation of which forms a culture of safety in children, responsibility for actions in the information space, strengthening spiritual and moral values, education of patriotism, the readiness of teachers and parents to accept the position of the child and respect for his independence [4].

III. METHODOLOGY

Before starting to write the article, various similar software tools should be researched. Their characteristics, demand and ease of use were analyzed.

Nowadays, there is various freely available parental control software, such as: Family GPS tracker KidControl, Kids Place, Kidslox, Kids Zone, Screen Time, Family Link [5] and many others. Among all these programs, we can single out the Microsoft Family Safety program, the use of which

will help monitor the activities of the child on their devices (phone, tablet, or computer), which run the Windows operating system without much effort.

An operating system is a set of programs that manage the resources of a computer, laptop, or any other gadget. According to statistics (see [9]), the use of the Android operating system is advanced for Ukraine in the mobile sector, while Windows is recognized as a desktop platform.

Table 1. Analysis of popular software for monitoring time spent on mobile devices and computers.

	<i>Kids Place</i>	<i>Kidslox</i>	<i>Screen Time</i>	<i>Family Link</i>	<i>Microsoft Family Safety</i>
Support operating system Windows	-	+	-	-	+
Support operating system Android	+	-	+	+	+
Content filtering	+	+	+	+	+
Program management	+	+	+	+	+
Control the time of use of the device	+	+	+	+	+
Geolocation	+	+	+	+	+
Lock device	-	-	+	+	+
Localization (Ukrainian)	-	-	-	+	+
Method of distribution	free	free	free	free	free

Based on the analysis, it was concluded that the Family Link [5] and Microsoft Family Safety programs are the most appropriate to use. Let us consider in more detail the features of the use of the Microsoft Family Safety software.

IV. RESULTS

By following these steps, you can ensure your child's safety while using the network and solve many other problems. The child should know these measures, and adults need to prepare it to consider into account psychological and age characteristics. After all, methods of ensuring the information security of children and educational impact should be adequate to the age and level of development of the child, to get the optimum result of educational activities and ensure safe socialization of the person in the information environment [8].

The education system plays an important role in ensuring children's information security, because of this a culture and a system of competencies in the field of information security are formed. No less important is the communication with parents to ensure the information security of children. There are specialized web resources and applications, the use of which allows even beginners to effectively use network resources and protect themselves and their child from unwanted content. It is the parents who must control this to prevent the undesirable effects of the overuse of certain devices.

In general, parental control applications are set to address two main tasks. First, create a safe play area for the child and protect the device from the chaotic effects the child. Second, control the list of programs, games, and the time of their use by the child.

Software tool Microsoft Family Safety from the company Microsoft – a free service that provides parental control (monitoring) devices running Windows 10 (Figure 2), Xbox and Android devices (Figure 1), running Microsoft Family Safety for users with a Microsoft account. Using the service settings on the resource <https://account.microsoft.com/family> (Family.Microsoft.com), you can get reports on your child's actions, set screen restrictions, check your child's location, restrict access to selected websites and games, and set restrictions and confirm purchases of software products from the Microsoft Store.

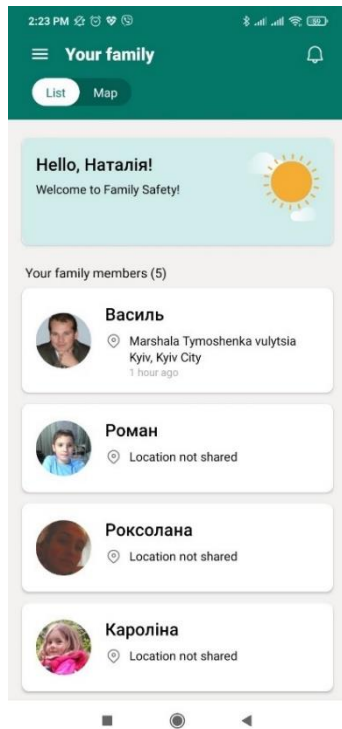


Figure 1. Manage accounts using a mobile application

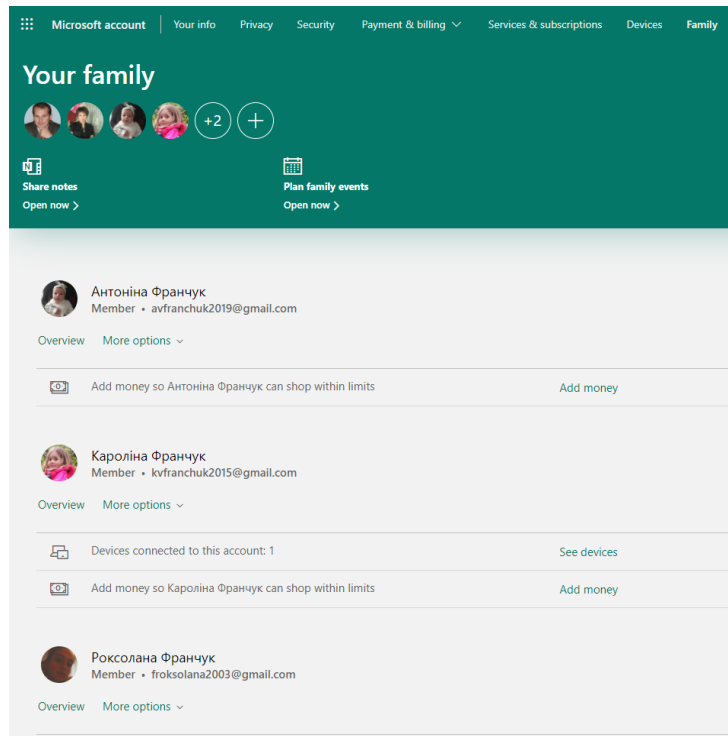


Figure 2. Manage accounts using a browser application

To use the Microsoft Family Safety software in the family, each member of the family must first have a Microsoft account. When you first log in to Family.Microsoft.com, you need to create a group for one of the family members (father or mother) using the Create Family resource. Then follow the instructions (see [10]) you can invite participants to join the Family group and start customizing controls. You can also use family groups on Xbox to add family members [3].

You can manage accounts using a mobile software tool (Figure 1) or a browser application (Figure 2).

There are many benefits to using Microsoft Family Safety software, such as:

- *Receiving reports on action (Activity)*. This email receives weekly reports of your child's online activity on Windows 10 and Xbox devices, as well as on Android devices with Microsoft Family Safety installed. You can also view actions at family.microsoft.com at any time. You can see the websites that children visit, their browser logs, the programs, and games they use, and how long they have used a particular device. Action reporting is a great way to keep track of events and adjust family settings accordingly (Figure 3).

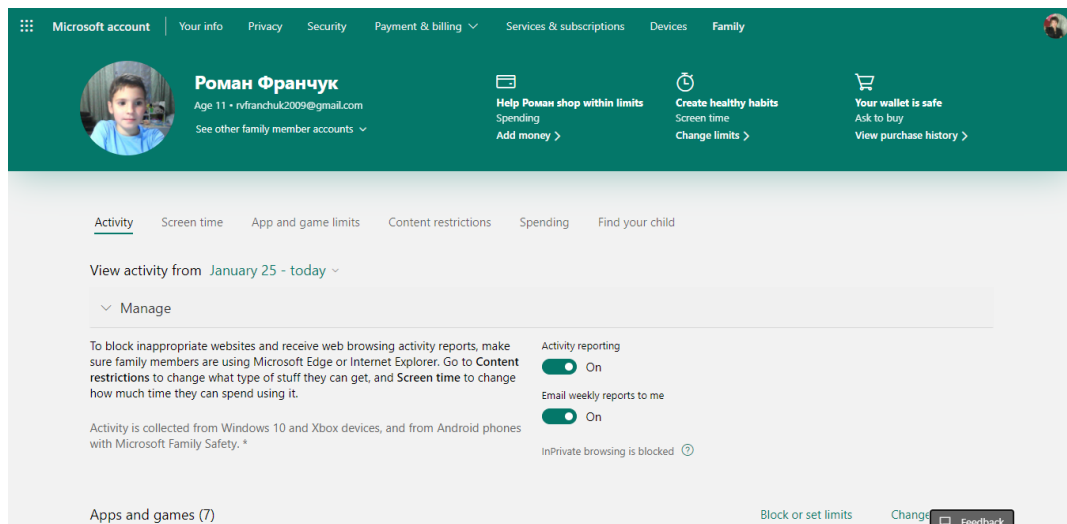


Figure 3. View user activity.

- *Planning for the time of the devices (Screen time)*. If desired, you can schedule a time when the child should use a particular device. This way you can be sure that your child has a good balance between the time spent in front of the screen and other activities. You can even set a time limit for certain programs or games (such as Minecraft). To do this, select a limit (device or application) and set the time. The child can ask for more time, then an e-mail arrives and can be easily answered, in the Microsoft Family Safety program or on Family.Microsoft.com (Figure 4).

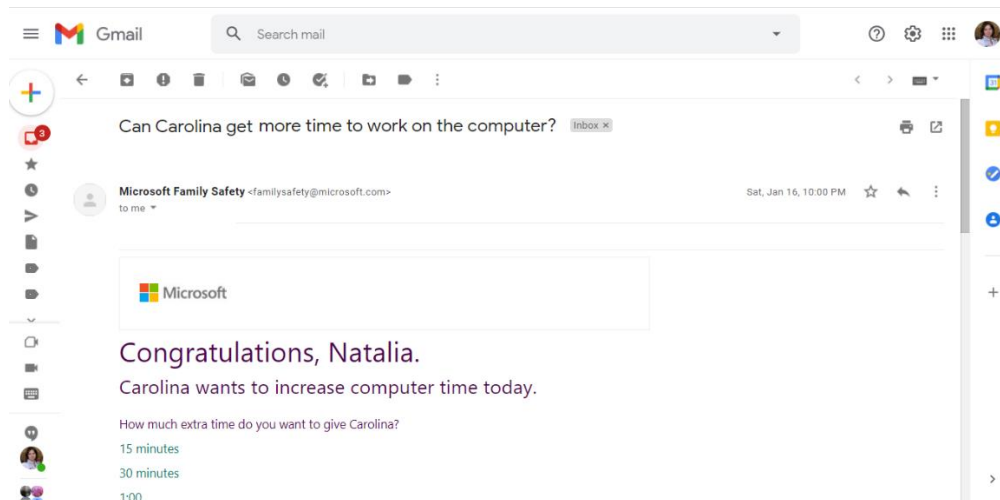


Figure 4. Request to extend the working hours.

- *Setting limits on the use of applications and games*. If a child spends too much time on certain programs or games, you can set limits on their use to achieve balance. Restrictions on apps and games are available on Windows 10 and Xbox devices. Once the restriction is set, it will apply to all the child's devices. (For example, giving a child two hours on YouTube, will consider the total time on YouTube with Windows 10, Xbox, and Android.) (Figure 5).

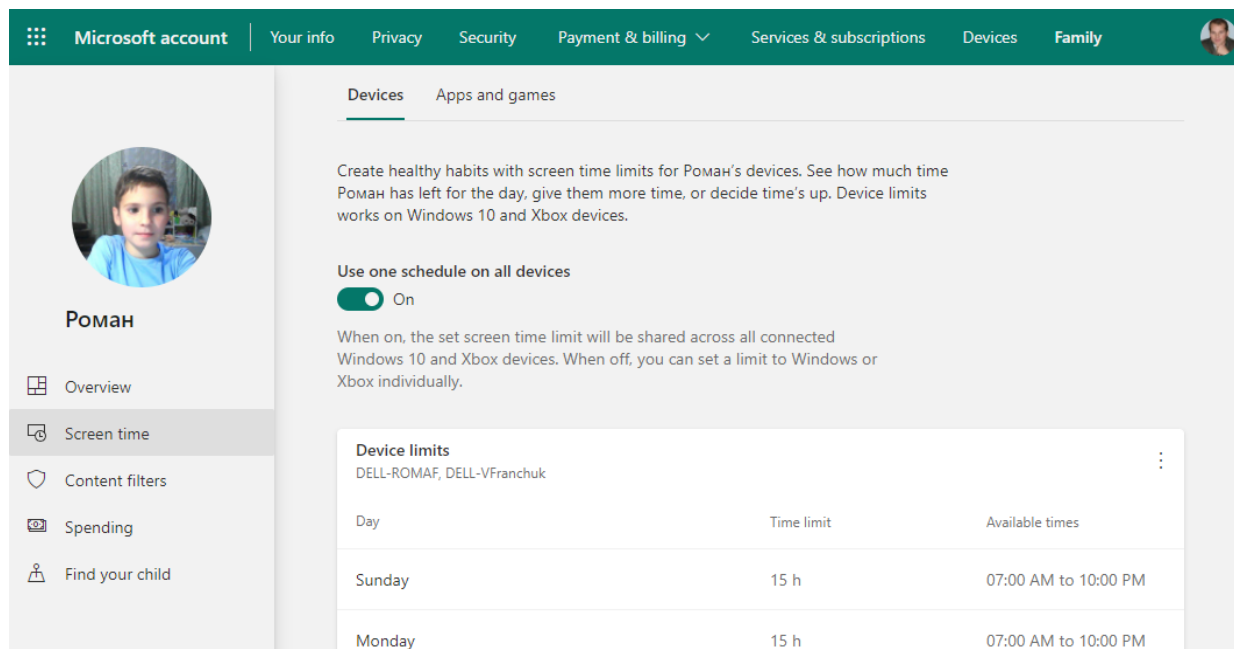


Figure 5. Time spent on devices.

- *Setting content filters (Content Filters)*. Content filters can be used to determine which websites a child visits and which games they can play. To do this, you need to set rules and decide under what conditions they can be violated. To set restrictions for children on the Internet, you can filter out inappropriate programs and games (set age limits for filtering) on devices running Windows 10, Xbox, and Android. Also, after setting age limits, many websites are blocked automatically. Exceeding the age limit will require an adult's permission. The child can send requests to view the

content. These requests will be sent to you by e-mail. They can be approved or rejected in the Microsoft Family Safety web application or on Family.Microsoft.com (Figure 6).

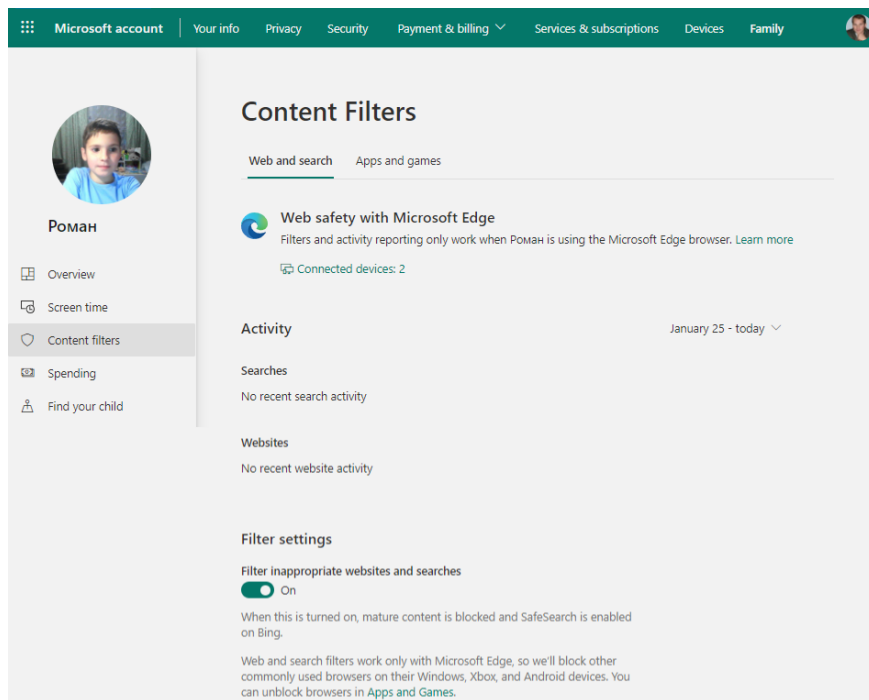


Figure 6. Setting content filters

- *Cost management of the child (Spending).* You can view your child's recent purchases, add funds to your Microsoft account, and decide if users need to be able to purchase something from the Microsoft Store (Turn on "Ask to buy") (Figure 7).

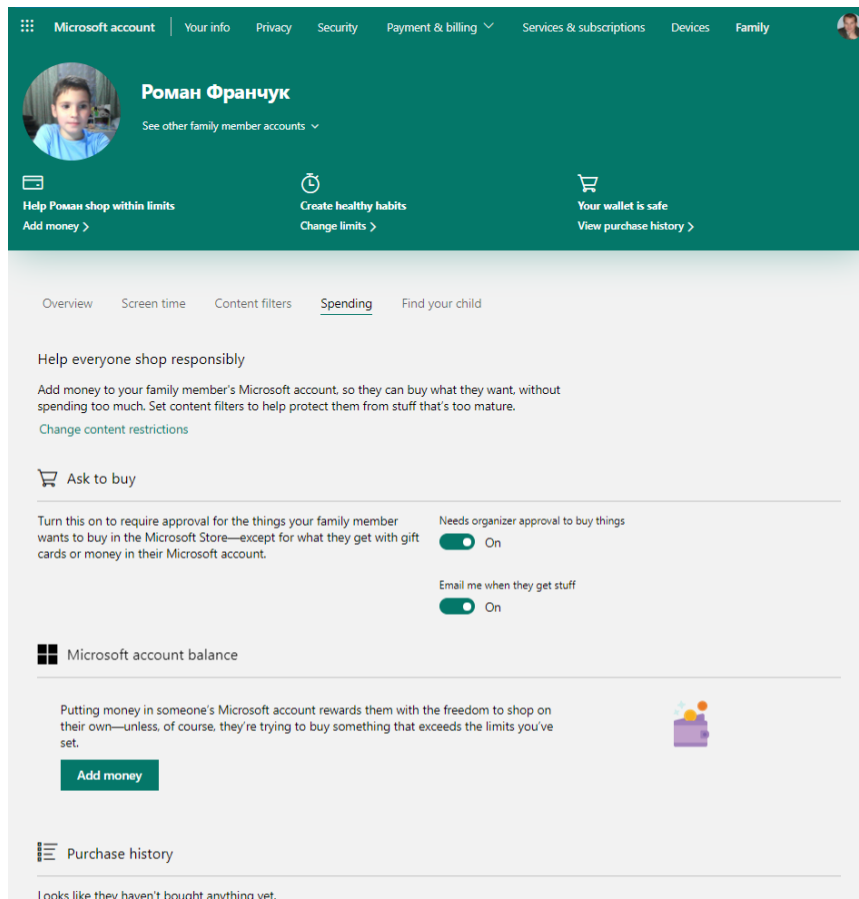


Figure 7. Cost management child

- *View the location of the family on the map (Find your child).* You can view where family members are at any time. Make sure your child gets home safely after school or review their location when she is with friends. To do this, open Microsoft Family Safety and look at the shared family map (see Figure 8).

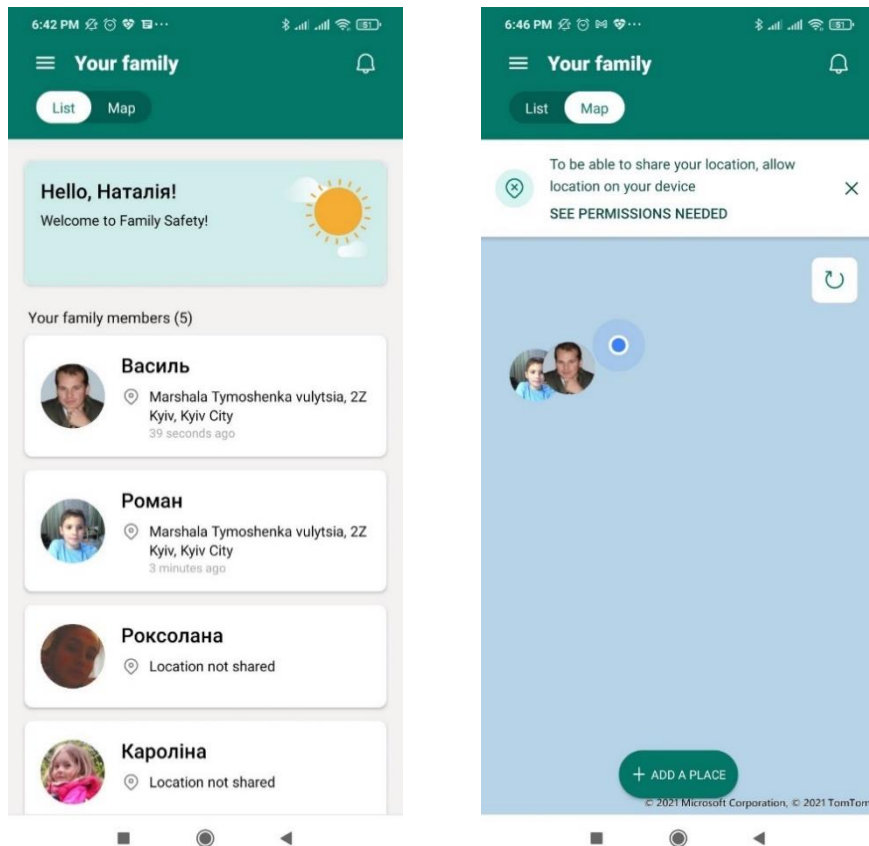


Figure 8. View the location of the family on the map.

When using Microsoft Family Safety, you must first consider the need to help children learn technology safely. By the way, this is one of the main goals of Microsoft. In particular, in many countries around the world there are various provisions: the law on the protection of children's privacy on the Internet (COPPA – USA), the General Data Protection Regulation (GDPR – European Union), the law on personal data protection (PIPA – South Korea).

V. DISCUSSION

This study examines one of the aspects of children's information security, namely monitoring children's work on mobile and personal devices. This study can be used by teachers and parents for the development of information culture.

Using Microsoft Family Safety provides parents with tools to monitor their child's actions on devices and the Internet, but this does not make the World Wide Web safer. The main task of using the application is to help parents choose child-friendly content and explain to them how to use their devices and Internet resources.

VI. FUTURE RESEARCH

In further research, we plan to pay attention to the vulnerability of such software programs. And to study the negative and positive impact, as well as psychological state of children when using parental control software.

VII. CONCLUSIONS

Therefore, the problem of the overuse of mobile devices requires constant analysis and coverage. It is clear, that the development, creation, and use of software to monitor the stay of children on the devices is one of the urgent needs of parents and teachers. The use of such software is just one of the tools to prevent the "collage of modern thinking" of children.

Solving these problems requires a comprehensive solution for the use of devices in the educational process, and teachers, in addition to teaching students, need to draw parents' attention to the use of mobile and stationary devices by children outside the educational institution. One of the tools for such parental influence can be the Microsoft Family Safety software.

VIII. ACKNOWLEDGMENT

Authors express gratitude to the developers of parental control software programs. They are also grateful to their four children who helped to carry out this research. It is worth noting that parental control is established on all devices they use for learning, communication and entertainment.

IX. DISCLOSURES

The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- [1] S. S. Andrieieva, Learning gadgets. *Computer at school and family*, vol. 8, 2016, pp. 53-55
- [2] Create a Google Account for a child. Accessed: Jan. 20, 2021. [Online]. Available: https://support.google.com/families/answer/7103338?hl=uk&ref_topic=7336731
- [3] Family & online safety. Accessed: Jan. 25, 2021. [Online]. Available: <https://support.xbox.com/en-US/help/family-online-safety/browse>
- [4] N. P. Franchuk, O. Yu. Rokytska, Information security of children on the Internet *Problems of informatization of the educational process in school and higher pedagogical educational institution* proceedings of the All-Ukrainian scientific-practical conference October 10 Kyiv: NPU imeni M.P. Dragomanova, 2017, pp. 142-143
- [5] V. M. Franchuk, N. P. Franchuk. Use of Family Link by parents and children *Computer at school and family*, vol. 1, 2020, pp. 34-39
- [6] How to enable parental controls on your existing Google Account. Accessed: Apr. 23, 2022. [Online]. Available: https://support.google.com/families/answer/9055704?hl=uk&ref_topic=7336731
- [7] H. M. Hych, "Clip" thinking of young people: friend or foe of learning? Scientific works [Black Sea state. Petro Mohyla University of the Kyiv-Mohyla Academy complex] Series: Pedagogy T. 269, Output 257, 2016, pp. 38-42
- [8] V. N. Kovalchuk, Problems of information security of children of different age categories *Computer at school and family*, vol. 8, 2010, pp. 58-62
- [9] The world's most popular operating systems 2020 Marketer. Accessed: Nov. 30, 2020. [Online]. Available: <https://marketer.ua/ua/stats-operating-system-2020>
- [10] What is a Microsoft family group? Accessed: Sep. 30, 2022. [Online]. Available: <https://support.microsoft.com/en-us/account-billing/what-is-a-microsoft-family-group-b6280c9d-38d7-82ff-0e4f-a6cb7e659344>

Про сучасне різноманіття технологій та архітектурних рішень Інтернету речей

Олена Смалько

*Кафедра комп'ютерних наук
Кам'янець-Подільський національний університет імені Івана Огієнка
Кам'янець-Подільський, Україна*

Анотація—У роботі розповідається про сучасне різноманіття технологій та архітектурних рішень Інтернету речей, обґрунтовується необхідність вивчення IoT-концепції в закладах вищої освіти, описується як найбільш доцільно конструювати відповідні навчальні курси та чим саме варто наповнювати використовувані в освітньому процесі навчально-методичні матеріали.

Ключові слова—*Інтернет речей, IoT-система, IoT-мережа.*

I. ВСТУП

Під «Інтернетом речей» (англ. Internet of Things, скорочено IoT) прийнято розуміти концепцію мережі, яка складається з пристроїв, оснащених сенсорами, датчиками та засобами передавання сигналів, а також програмними застосунками, що забезпечують обмін даними з метою віддаленого моніторингу за станом об'єктів, аналізу зібраної інформації та прийняття на її основі рішень. У наш час IoT-мережі набувають значного поширення, їх все частіше використовують для підтримки ведення домашнього господарства, у розвитку інфраструктури територіальних громад, в комерційній та виробничій діяльності, для покращення якості повсякденного життя людей та оптимізації природокористування. Тому дуже важливо, зокрема під час формування цифрових компетентностей здобувачів освіти та всіх зацікавлених у цьому людей, навчати їх розумінню нових можливостей, що відкриваються з використанням IoT-систем у різних галузях діяльності людини.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

З початку 2000-х років активно розвиваються технології Інтернету речей, постійно вдосконалюється та урізноманітнюється використовуване апаратне забезпечення, зростає кількість спеціалізованих програмних і мобільних застосунків, що допомагають фахівцям аналізувати дані, отримані від IoT-пристроїв. Про весь цей корисний інструментарій можна знайти безліч різних публікацій, які знаходяться в тому числі у відкритому доступі. Останнім часом почали з'являтися і фундаментальні наукові праці [2], а також окремі навчальні видання [1], орієнтовані на україномовного читача, котрий цікавиться концепцією IoT, в яких на різному рівні подаються основи функціонування та використання сучасного Інтернету речей.

У даній публікації описується авторське дослідження [3] та заплановані шляхи його продовження, які закладатимуть фундамент для більш ґрунтовної практичної роботи студентів, орієнтованої на проектування IoT-рішень різного призначення в рамках однойменного університетського навчального курсу.

III. МЕТОДИ

Автором проведено детальне теоретичне дослідження передумов формування концепції Інтернету речей, проаналізовано різні моделі архітектурних рішень IoT та існуюче різноманіття технологій, що застосовуються на фізичному та мережевому рівнях. Значну

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

увагу приділено вивченню використовуваних в Інтернеті речей протоколів передавання даних, особливостей граничних і туманних обчислень, що є розширеннями хмарних обчислень. Систематизовано відомості про популярні варіанти живлення IoT-систем, а також про нетрадиційні підходи до збирання енергії. Досліджено основи проектування Інтернету речей, приклади його реалізації в різних сферах життєдіяльності людей, зроблено огляд різноманіття IoT-платформ, проміжного програмного забезпечення та застосунків, що допомагають кінцевим користувачам отримувати користь від використання Інтернету речей.

IV. РЕЗУЛЬТАТИ

За результатами проведеного дослідження підготовлено навчально-методичні матеріали, в яких описано типи широко використовуваних мікроелектромеханічних систем, датчиків та інші категорії граничних пристроїв Інтернету речей, подано важливу інформацію стосовно надійних та енергоефективних методів передавання даних в IoT-системах; роз'яснено особливості різних способів генерування енергії для пристроїв IoT, що стали популярними серед користувачів завдяки наукоємності використаних при їх створенні технологій, безпечності, зручності, вигідної ціни та тривалої роботи відповідного обладнання. Також розроблено серію завдань по дослідженню можливостей використання IoT в бізнесі, промисловості, в ритейлі, у транспортній логістиці; по вивченню моделей Інтернету речей, які можуть бути найбільш доцільними для забезпечення критично важливих комунікацій (на рівні підприємств і в територіальних громадах), зокрема для належного вирішення проблем громадської безпеки, прискореного реагування на надзвичайні ситуації; по ознайомленню з перспективами застосування IoT в галузі охорони здоров'я, персоналізованої медицини, у фармацевтичному виробництві, в обороні та військовій сфері.

З метою підготовки до розв'язання практичних задач досліджувались можливості адміністративної панелі для збирання та обробки даних про споживання електроенергії, води, газу й тепла Jooby RDC Dashboard (сумісної україно-швейцарської розробки), окремих IoT-рішень ізраїльсько-української компанії Eastern Peak (інтелектуальної системи управління теплицею Green IQ, а також мобільних застосунків Cobra Connect для дистанційного керування моніторингу багатьох функцій автомобіля та DOORe – системи безпеки домашнього входу з розумним замком). Вивчались функції смарт-динаміка Google Home, GSM сигналізації SafeHome, а також досвід роботи української компанії Intellias по розробці екосистеми для смартизації сільськогосподарських технологій.

V. ОБГОВОРЕННЯ

Для якісного навчання в закладах вищої освіти є вельми вагомою грамотно продумана теоретична підтримка освітнього процесу. Саме тому дуже помічними для викладачів і здобувачів освіти вважаються ті навчальні видання, що містять добре систематизовану інформацію, яка ще й збалансована у відповідності з навчальним навантаженням. Добираються та компонуються навчальні матеріали в таких виданнях по-різному, залежно від найменування навчальної дисципліни, від того, які знання та навички покликана формувати у студентів, та яку саме спеціальність вони здобувають.

Розгорнені інформаційні матеріали, обґрунтовані методичні вказівки та якісна наочність, що конче потрібна для підтримки вивчення курсу, покликано формувати у суб'єктів навчання навички проектування в середовищі Інтернету речей, традиційно мають містити історичні відомості, спираючись на які можна цілком аргументовано переконувати студентів у важливості аналітичної діяльності в прогресивному розвитку людства та в неоціненності ролі окремих особистостей у покращенні нашого життя. Це може спонукати зацікавлених молодих людей до подальшої ефективної навчальної діяльності та значних звершень у майбутній професії.

До того ж, не зважаючи на цілком резонну вимогу, що здебільшого висувається до навчальних видань стосовно неприпустимості перевантаження суб'єктів навчання надмірним обсягом інформації, поданої в них, важливо подавати релевантну, ретельно структуровану і в

той же час детальну інформацію, якої буде достатньо для успішного засвоєння відповідного курсу.

Якщо мова йде про навчально-методичні видання, то в них ще обов'язково подаються методичні основи викладання навчальної дисципліни. Бажаними є матеріали практичного змісту. Так, наприклад, при дослідженні різноманітних галузей застосування Інтернету речей можна пропонувати завдання по вивченню поширених функцій сучасних розумних будинків, моделей і технологій розумних міст, включаючи аналіз популярних (у тому числі вітчизняних) IoT-рішень, призначених для домашньої автоматизації, для розумного обліку ресурсів житлово-комунального господарства, розумного вуличного освітлення, розумного транспортного забезпечення, інтелектуальних систем паркування тощо.

Потрібно ознайомлюватись і з перевагами, які надають системи Інтернету речей у плануванні ресурсів підприємств, у підвищенні ефективності бізнесу завдяки автоматизації процесів виробничого циклу підприємств, у моніторингу ланцюга поставок товарів і послуг. Корисним буде й вивчення особливостей керування даними в сільськогосподарських екосистемах для ефективного ведення агробізнесу, для дистанційного моніторингу стану тваринницьких ферм.

Також важливо давати студентам завдання по проектуванню власних моделей систем Інтернету речей бізнесу та «розумного» облаштування свого будинку із зазначенням конкретних технічних і програмних рішень.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У подальшому планується розробити завдання для лабораторних занять, під час яких будуть використовуватись апаратно-програмні платформи на основі мікрокомп'ютерів Raspberry Pi, мікроконтролерів Arduino тощо [4].

VII. ВИСНОВКИ

Сучасні технології Інтернету речей все активніше впроваджуються в наше повсякденне життя, в бізнес, в культуру природокористування. Вони покращують побут, допомагають економити час і кошти, турбуватись про здоров'я та екологію. Всепроникний Інтернет речей беззаперечно змінює наше сьогодення, і вже очевидно, що без опанування навичок керування ним нам не обійтись у майбутньому. Тому необхідно цілеспрямовано формувати ці навички у здобувачів освіти, які навчаються на різних освітніх рівнях, особливо у студентів-фізиків та майбутніх IT-фахівців.

ЛІТЕРАТУРА

- [1] Б. Ю. Жураковський та І. О. Зенів, *Технології інтернету речей. Навчальний посібник*. Київ: КПІ ім. Ігоря Сікорського, 2021. Режим доступу: <https://ela.kpi.ua/handle/123456789/42078>. Дата звернення: Жовт. 30, 2022.
- [2] О. А. Баранов, *Інтернет речей: теоретико-методологічні основи правового регулювання*. Т.1: Сфери застосування, ризики і бар'єри, проблеми правового регулювання: монографія. Київ: Видавничий дім «АртЕк», 2018.
- [3] О. А. Смалько, Укл., *Використання Інтернету речей: навчально-методичний посібник*. Кам'янець-Подільський: Друкарня «Рута», 2022.
- [4] 20 інноваційних та найкращих програм IoT для підвищення вашого потенціалу IoT. Режим доступу: <https://ciksiti.com/uk/chapters/6191-the-20-innovative-and-best-iot-software-to-boost-your-iot-po>. Дата звернення: Жовт. 30, 2022.
- [5] IoT Architecture: the Pathway from Physical Signals to Business Decisions. Available: <https://www.altexsoft.com/blog/iot-architecture-layers-components>. Accessed on: Oct. 30, 2022.

Informational system of relief generation based on Perlin's noise mathematical algorithm

Vyacheslav Fomenko, Oksana Shovkoplias

*Department of Computer Science
Sumy State University
Sumy, Ukraine*

Abstract— When solving some complex physical problems, for example, calculation the trajectory of the movement of a barrel or a wheel rolling down a mountain needs to be generated the surrounding area. And in this article the most popular method for generation surrounding area described.

Keywords— *Generation algorithms, Terrain generation, Computer graphic, Procedural generation, Perlin Noise.*

I. INTRODUCTION

Nowadays days everyone was striving for photorealism, but it was always lacking. Generation algorithms was developed to overcome this problem the main point of this algorithms is to generate realistic graphic using minimal resources for that, and as the result the generated content should be differ from the content that was made by a human.

II. RELATED WORK

In 2013, Genevaux et al. modified the algorithm and created a new way of generating the landscape based on hydrology. Their algorithm follows the input contours of the terrain and the river on it, given by the user, the algorithm will take the input information to generate a complete drainage network, which leads to the formation of rivers that flow from the sources to the sources, with the expected variation in height. After the rivers are created, they start working with the terrain, modifying the terrain according to the river so that the overall result of the map is topologically correct.

III. METHODOLOGY

To begin with, it is necessary to define what procedural generation is and what digital content generation is. The generation of digital content is a subset of algorithms, it is a key component of modern methods of creating films, video games, etc., which is formed by a set of algorithms for creating landscapes, maps, characters, weapons, various effects: fire, fog, rain, etc., ensuring the reduction of production costs.

Procedural generation – automatic creation of content using algorithms. In other words, procedural generation is software that can independently create unique content. Under content in this case, it is the generation of digital content that was given above. The key feature of the created content is that it should be valid, that is, if it is a game, then the generated content should not prevent the player from passing the level, if it is a movie, then it should look natural, etc. A good random number generator produces numbers that are unrelated and show no discernible pattern. As we're beginning to see, a little randomness can be a useful thing when programming organic, realistic behavior. However, randomness as the sole guiding principle is not necessarily natural [1].

There is an algorithm that produces more natural results known as Perlin noise. Perlin noise was developed by Ken Perlin he invented it while he worked on the original Tron movie in 1980s; it was used it to create procedural textures for effects. He won an Academy Award in 1997, for technical
2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

achievement for this work. It can be used to create a variety of effects with natural qualities such as clouds, landscapes and patterned textures such as marble.

Perlin noise has a more organic appearance because it produces a naturally ordered ("smooth") sequence of pseudorandom numbers. The graph below shows Perlin noise over time, where the abscissa represents time; note the smoothness of the curve.

Noise is a random number generator in computer graphics. It's a random, unstructured pattern that's useful in those cases where you need a source of extra detail that's missing from the obvious structure the example of noise function is represented on [2, Fig. 1(a)].

Perlin noise is a multidimensional algorithm used in procedural generation, textures, terrain generation, map generation, surface generation, vertex generation, and so on.

We assign a number from 0.0 to 1.0 to each point on the card. In this image, 0.0 is black and 1.0 is white. Noise itself is just a set of numbers. We need to give it meaning. The first thing you can think of is to bind the noise value to the height (this is called a height map) [3].

This algorithm is good for creating different procedural textures but also has some disadvantages. Perlin's noise is noticeably quadratically shifted. It is smooth and random as the formulas require, but the features and shapes that the noise creates are mostly aligned at 45 or 90 degrees to the input coordinates. This phenomenon persists globally throughout the noise, offering a small break or gap between them. The Perlin noise algorithm is based on a square, cubic, or in general hypercubic grid. At each vertex of this grid, it selects tilt directions (gradient vectors), calculates their extrapolations, and then blends them together with their neighbors. Its displacement is a direct consequence of this grid structure and how the gradients interact on it. New gradients appear only at the vertices of the grid, and their protrusions merge only through the vertices of the same cell. Objects with better mesh angles merge effortlessly, while those with other angles are often not as successful [4].

IV. RESULTS

As a result of work for this algorithm we should get a realistic relief based on input parameters that could be used for future projects the result is represented on the **Fig. 1(b)**.

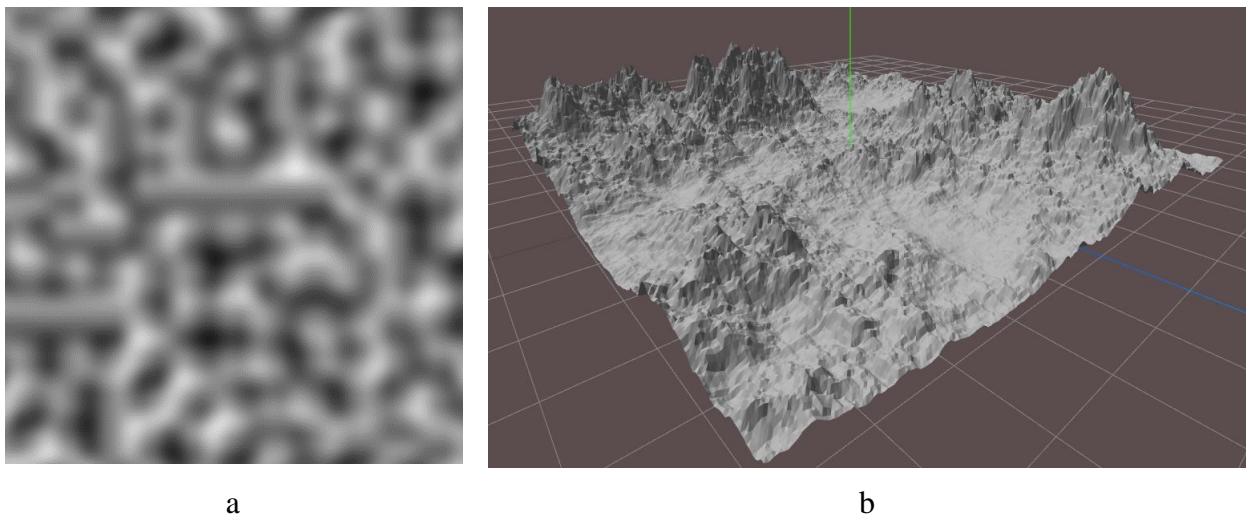


Fig. 1 – Example of noise function (a), Algorithm work result (b)

V. DISCUSSION

This work will help to make procedure of generation more easier and will help to avoid some disadvantages and make it work a little bit faster of Perlin noise. But also it will not help to overcome the main problem that Perlin's noise is noticeably quadratically shifted. It is smooth and random as the formulas require, but the features and shapes that the noise creates are mostly aligned at 45 or 90 degrees to the input coordinates. This phenomenon persists globally throughout the noise, offering a small break or gap between them.

VI. FUTURE RESEARCH

In future this method could be improved to overcome problems that this method has and to make it more detailed.

VII. CONCLUSION

Perlin noise algorithm was described in this article to show how we could generate random computer graphics objects and effects also his advantages and disadvantages was shown. As a result Perlin noise is good algorithm for generating realistic computer graphic that don't need a lot of resources.

REFERENCES

- [1] *Jessica Van Brummelen and Bryan Chen* "Procedural Generation." https://www.mit.edu/~jessicav/6.S198/Blog_Post/ProceduralGeneration.html (accessed Oct. 20, 2022)
- [2] *Amit Patel* "The Perlin Problem"<https://noiseposti.ng/posts/2022-01-16-The-Perlin-Problem-Moving-Past-Square-Noise.html> (accessed Oct. 20, 2022)
- [3] "An Introduction to Fractal Image Compression Literature Number: BPRA065 Texas Instruments Europe," 1997
- [4] *J. D. Génévaux and B. Beneš*, "Terrain generation using procedural models based on hydrology," vol. 32, *ACM Trans Graph*, no. 4, Jul. 2013

Software Architecture Design of Software Developed in Student Engineering Teams

Ihor Polataiko

*Department of Information Technology
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine*

Abstract—This paper presents the process of architecture design of the software, intended for development in student engineering teams. Software architecture is a primary driver of software quality. Considering the fact, that students with little experience are working on software, that is intended for real-world usage in the educational institution, it is required to define a deliberate process of software architecture design. The process, developed as the result of the performed research, defines the main stages in architecture design, describes the set of related documentation, and preferable ways of decision-making and validation.

Keywords—*Software Engineering, Software Architecture, Process, Education, Student Teams.*

I. INTRODUCTION

To grow successful software engineers, educational institutions, that teach software engineering-related fields should provide students not only with theoretical knowledge and practical assignments but also with a significant amount of production-close development experience. Having an environment, where students with very little or without industry experience could be able to practice and gain required skills is essential. Also, institutions, usually, have some category of students, which are typically in their senior years at the institution, who has some good amount of industry experience (1 or 2 years) and would like to teach and mentor more junior students. This type of experience could be incredibly useful for the mentioned category of students because it gives them the ability to better structure their own knowledge while teaching other students, and receive hands-on experience of technical leading of the projects. Apart from that, most educational institutions have an information technology (IT) department, that manages and develops software for internal needs. Students engineering organization, called PNUdev, was created and currently functions at Vasyl Stefanyk Precarpathian National University.

Software, developed by the teams, that form student software engineering organization should be of high enough quality. During the time of organization functioning, it was observed, that most of the significant reworks to the code changes, made by team members in form of merge requests, were due to design decisions. It was decided to define an upfront architecture process, which teams should follow for each project, that would help to reduce the number of reworks and increase the overall quality of the architectural design of the software along with increasing the level of awareness of the architectural decisions.

II. RELATED WORK

Multiple studies were conducted in the field of innovative approaches to software engineering education and project management processes. Paper [1] describes an application of short-term projects in cross-cultural environments for improving the soft and technical skills of the students. In papers [2], [3] researches related to increasing the level of student motivation and engagement in project-based learning are described. Papers [4], [5], focus on the development of real-world software by students as a vital aspect of engineering education. The research described in the paper [6] development management process, based on reflective weekly monitoring, for the context of 2022 International Conference on Innovative Solutions in Software Engineering Ivano-Frankivsk, Ukraine, November 29-30, 2022

student projects. Finally, studies [7] and [8] are related to using code review as a learning tool for software engineering students. Also, the studies [9], [10] were using parts of the software products, developed by PNUdev, as supportive tools.

III. METHODOLOGY

During the research, both the process of architecture decisions making and the architectural template document were developed as the optimal process and result of the application architecture definition phase in the given context of the student software engineering teams, working on the project. An architectural template helps to unify the form in which the decisions regarding the architecture and design are presented. Decisions are made in the group meetings, where all the team members are included.

The first phase of the architectural definition process is the definition of contexts. At this stage, features are grouped together into contexts. Contexts should have a minimal dependency and relations between each other. If the product is not new, contexts from the architectural template of the previous project are used as a base. In case, new contexts are identified, they should be marked as “new” in the architectural template. Contexts should not share domain data models but can reference aggregate roots from different contexts. For each new context, the decisions regarding dependencies on other contexts, level of splitting, and communication with other contexts should be made. The level of splitting can be either on code level: separate module or component, or deployment level: separate application. Communication can be either synchronous or asynchronous via messaging broker. Asynchronous communication is preferred. As the result of this phase, the diagram of contexts should be created which presents all the contexts, dependencies between them, and ways of communication. The diagram should include deployment boundaries as well. Also, sometimes, one context can be spitted across a couple of deployment units, e.g. back end and front end, which have the same contexts but run in different environments. An example of the most basic context diagram without dependencies between contexts is presented in figure 1.

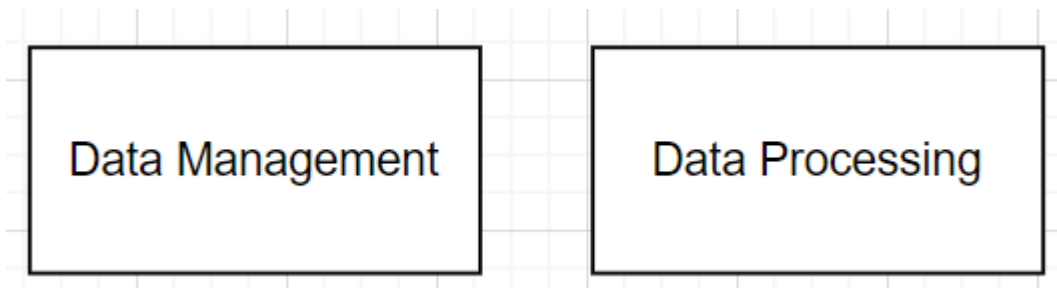


Figure 1. Example of context diagram without dependencies

After contexts are defined, there should be a process to define which of them might be partially or fully covered by the existing open source software systems in order to not reinvent the wheel. In this case, requirements might be updated to focus on integration with that software for the contexts, that are outsourced. Also, at this phase, doubts regarding the technical achievability of the requirements should be cleared out by means of building small proofs of concepts for each of them, which is assigned to some project team members. If too much non-clarity and questions arise during this phase, the project might be suspended and a software research project might be conducted with the team if the context doesn't imply strict timelines.

The next phase was identified as domain data model definition. For each of the contexts entities have to be identified, based on the requirements for the features. Nouns in the requirements often represent some domain entity. All the functional requirements should be analyzed one by one to identify most of the entity attributes needed (the ones, that were missed, will be added during development). The result of this phase should be a domain data model diagram where all domain entities, their attributes, relations, and fields by which they will be accessed, are represented. If the context existed in the previous project related to the product, then the domain data model diagram

should be the updated version of the diagram from the previous project. Figure 2 presents an example of a domain data model diagram for the context.

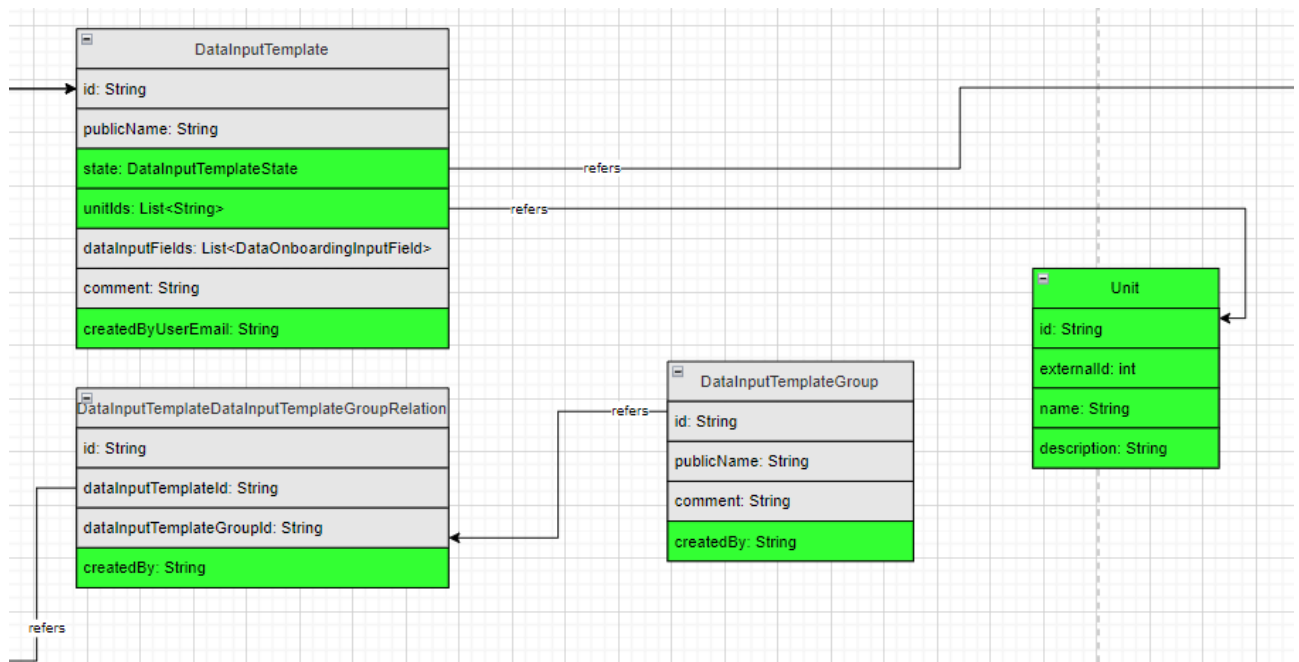


Figure 2. Example of a domain data model diagram

After the contexts and data model are defined, a detailed components diagram should be created. The diagram should specify applications created, show the communication of different components, integration with other systems, and specify the infrastructural components, such as databases, centralized caches, etc. At this phase, non-functional requirements have to be considered, when choosing different technologies and approaches. The decision regarding technical stack, usage of caches, serverless technologies, reactive vs non-reactive drivers, server-side rendering vs single page application should be made at this phase, based on the requirements and team experience. If the product is not new, the component diagram from the previous project should be used as a base. Figure 3 presents an example of a component diagram.

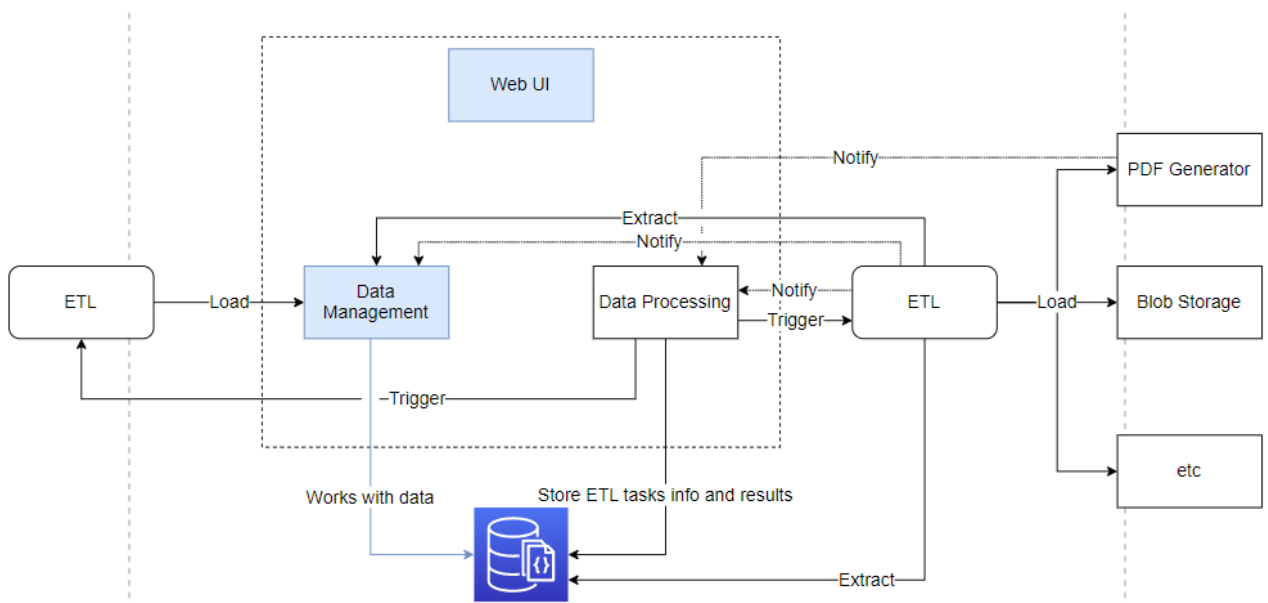


Figure 3. Example of a component diagram

For all parts of the project, which implies exposing the API either outside of the context or outside of the deployment unit, the APIs should be documented and added to the template with reference to the related feature. Apart from that, all the APIs used for integration with other

software systems should be also added to the architecture template.

The resulting architectural template should contain a context diagram, domain data model diagram for each context, component diagram, list of the APIs to be developed, and list of APIs for integration. It should be reviewed and validated carefully at a specific meeting, where all the involved people go through the filled architectural template for the project, discuss the questions and validate the architecture through the mental simulation of execution of functional requirements, having the APIs, components, and data models defined. Based on the results of the review, decisions, and the architectural template might be refined and reviewed again if needed.

IV. RESULTS

The research was performed by means of evaluating the experience of creation and leading of the PNUdev organization, experimentation, and path of trial and error. The organization functions for about 2 years, at the time of writing this paper. During this time, 5 software products were worked on. Some of them continue to evolve in the current projects. Since the start of applying the described template, the number of code review comments, related to the architectural design decisions, is reduced by 90%. The remaining 10% are related to human errors and lower-level design decisions.

V. DISCUSSION

A proposed process of software architecture design allows student engineering teams to have a straight and comprehensive process of defining architectural decisions upfront. Considering, that students, joining the engineering teams, usually, do not have enough expertise to be able to make good quality architectural decisions, the process helps the team to focus on development, during the development phase, and eliminates the possibility of making design decisions of low quality, because the primary aspects of the architecture for the project are defined and reviewed by the whole team and by more experienced colleagues, who manage the teams. Usually, the technical manager of the team, who has some industry experience, drives the meetings dedicated to architectural decisions makings. It is beneficial for both the technical manager and less experienced members of the team. Technical managers practice the skills of architectural decisions making while explaining the reasons for the decisions to the rest of the team, and other members of the team are able to follow the process of architectural decision-making and participate in it.

The model is not intended for application outside of the context of the student software engineering teams. Software engineers of a more senior level are able to make most of the decisions in the scope of the implementation. However, API-driven design, during which the list of APIs is negotiated and defined upfront, that model imposes is a good practice, independently of the context. Also, the described process assumes, that project development is made in medium-term iterations, hence projects, that practice continuous delivery of features to a production environment, would, most certainly, not be able to apply the described process, because it would be a huge overhead.

VI. FUTURE RESEARCH

The aspects, related to measuring and assuring the quality of architectural decision-making can be a focus of future research on the given topic.

VII. CONCLUSION

As the result of the performed research and analysis of the experience of software design and development performed by the teams which are part of the PNUdev organization, which functions at Vasyl Stefanyk Precarpathian National University, the process of architecture design and related set of documentation, called architecture template, was developed.

VIII. ACKNOWLEDGMENT

The research has been performed at and supported by the Department of Information Technology of the Vasyl Stefanyk Precarpathian National University.

IX. DISCLOSURES

The author of this paper is a student of the Vasyl Stefanyk Precarpathian National University at the time, of the research and writing this paper.

REFERENCES

- [1] I. H. Akerlund, G. Audemard, H. Bollaert, V. Hayenne-Cuvillon, A. Hlobaz, M. Kozlenko, P. Milczarski, J.C. Monteiro, J. Morais, D. O'Reilly, P. Possemiers, and Z. Stawska, "Project GGULIVRR: generic game for ubiquitous learning in interactive virtual and real realities," in EDULEARN20 Proceedings of the 12th International Conference on Education and New Learning Technologies, July 6-7, 2020, pp. 5973-5979, doi:10.21125/edulearn.2020.1566.
- [2] P. Morais, M. J. Ferreira and B. Veloso, "Improving Student Engagement With Project-Based Learning: A Case Study in Software Engineering," in IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, vol. 16, no. 1, pp. 21-28, Feb. 2021, doi: 10.1109/RITA.2021.3052677.
- [3] R. K. Pucher, A. Mense, H. Wahl and F. Schmöllebeck, "Intrinsic motivation of students in project based learning," in Transactions of the South African Institute of Electrical Engineers, vol. 94, no. 3, pp. 6-9, Sept. 2003.
- [4] R. J. Machado, P. Guerreiro, E. Johnston, M. Delimar and M. A. Brito, "Work in progress - IEEEExtreme: From a student competition to the promotion of real-world programming education," 2009 39th IEEE Frontiers in Education Conference, 2009, pp. 1-2, doi: 10.1109/FIE.2009.5350540.
- [5] Buhari, S. Valloo and H. Hashim, "A Streamlined Approach to Enhance the Capacity of Undergraduate IT Students to Deliver High Quality and Demand-Driven Final Year Project: A Conceptual Framework on Collaboration between Industry and University," 2017 7th World Engineering Education Forum (WEEF), 2017, pp. 910-914, doi: 10.1109/WEEF.2017.8467126.
- [6] G. M. Marques, S. F. Ochoa, M. C. Bastarrica and F. J. Gutierrez, "Enhancing the Student Learning Experience in Software Engineering Project Courses," in IEEE Transactions on Education, vol. 61, no. 1, pp. 63-73, Feb 2018, doi: 10.1109/TE.2017.2742989.
- [7] Z. Kubincová and I. Csicsolová, "Code Review in High School Programming," 2018 17th International Conference on Information Technology Based Higher Education and Training (ITHET), 2018, pp. 1-4, doi: 10.1109/ITHET.2018.8424617.
- [8] T. Brown, M. R. Narasareddygar, M. Singh and G. Walia, "Using Peer Code Review to Support Pedagogy in an Introductory Computer Programming Course," 2019 IEEE Frontiers in Education Conference (FIE), 2019, pp. 1-7, doi: 10.1109/FIE43999.2019.9028509.
- [9] Lazarovych et al., "Software Implemented Enhanced Efficiency BPSK Demodulator Based on Perceptron Model with Randomization," 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2021, pp. 221-225, doi: 10.1109/UKRCON53503.2021.9575458.
- [10] M. Kozlenko, O. Zamikhovska, V. Tkachuk and L. Zamikhovskyi, "Deep Learning Based Fault Detection of Natural Gas Pumping Unit," 2021 IEEE 12th International Conference on Electronics and Information Technologies (ELIT), 2021, pp. 71-75, doi: 10.1109/ELIT53502.2021.9501066.

Реалізація еволюційного алгоритму на реальному квантовому комп'ютері

Валерій Ткачук, Олександр Кулик

*Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація — Запропоновано та практично реалізовано новий класично-квантовий генетичний алгоритм (CQGA), орієнтований на виконання на реальному квантовому комп'ютері. Обмежена розрядність таких комп'ютерів не дозволяє реалізувати наближені алгоритми квантового клонування, тому для організації еволюційного процесу пропонується використовувати технологію томографії квантових станів. Можливості запропонованого підходу проілюстровано з використанням квантового комп'ютера IBM Q, для реалізації доступу до якого використано програмне забезпечення з відкритим кодом Qiskit.

Ключові слова — *Квантовий комп'ютер, Еволюційний алгоритм, Кубіт, Томографія квантових станів, Оператор квантового гейту.*

I. ВСТУП

Евристичні методи оптимізації, в основу роботи яких покладено ідеї квантових обчислень, орієнтовані на їх практичну реалізацію засобами класичних обчислювальних систем [1]. При цьому практично всі квантові оператори, та і в цілому сам квантовий генетичний алгоритм без принципових переробок не можуть бути реалізовані на квантовому комп'ютері. Успіхи у розробці квантових комп'ютерів та можливості вільного доступу до них дозволяють адаптувати та апробувати такі алгоритми для їх виконання хоча-б частково на реальному квантовому комп'ютері. Для ілюстрації можливості запропонованого алгоритму використано квантовий комп'ютер IBM Q із простим та зручним інтерфейсом для його програмування.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Квантові генетичні алгоритми ґрунтуються на поєднанні технології квантових обчислень із ідеями класичних генетичних алгоритмів. Однак розробка та дослідження цього класу алгоритмів завжди була орієнтована на ефективне розв'язання задач оптимізації без розгляду можливості їх реалізації на квантовому комп'ютері [2]. У зв'язку із значними успіхами у розробці як ефективних он-лайн симуляторів, так і доступних з використанням хмарних технологій квантових комп'ютерів за останні декілька років появилася незначна кількість публікацій, присвячених цій тематиці (див. наприклад [3]).

III. МЕТОДИ

На сьогодні з використанням хмарних технологій загальнодоступними є як 5-ти кубітні квантові комп'ютери компанії IBM (ibmq_santiago, ibmq_manila, ibmq_bogota, ibmq_quiti, ibmq_belem), так і ряд квантових онлайн-симуляторів (ibmq_qasm_simulator (32 кубіти), simulator_mps (100 кубітів)) [4]. Для доступу до них розроблено проект із відкритим кодом Qiskit, написаний на Python, що включає в себе набір інструментів для написання програм на рівні квантових схем.

У зв'язку із неможливістю клонування невідомих квантових станів [5] для організації еволюції популяції в часі пропонується гібридна, класично-квантова структура генетичного квантового алгоритму:

1.	$t \leftarrow 0$
2.	ініціалізація $Q(t)$
3.	томографія $Q(t)$ та перехід до $P(t)$
4.	оцінка пристосованості $P(t)$ та пошук найкращої особини популяції
5.	while (умова завершення еволюції)
6.	$t \leftarrow t + 1$
7.	побудова наступного покоління $Q(t)$
8.	томографія $Q(t)$ та перехід до $P(t)$
9.	оцінка пристосованості $P(t)$ та пошук найкращої особини популяції
10.	end while

тут: $Q(t)$ – квантовий реєстр популяції розв'язків на момент часу t ; $P(t)$ – класичний реєстр представлення популяції, отриманий в результаті томографії квантових станів. Важливим у пропонованому алгоритмі є те, всі еволюційні оператори, такі як створення, еволюція та мутація в популяції реалізовано з використанням квантового комп'ютера.

IV. РЕЗУЛЬТАТИ

У Для ініціалізації початкової популяції $Q(0)$ використано генератор випадкових чисел, реалізований на самому квантовому комп'ютері [6]. Початковий кут θ_i для кожної особини популяції розраховується відповідно до квантового протоколу, представленому на рис.1 (для 5-кубітного квантового комп'ютера):

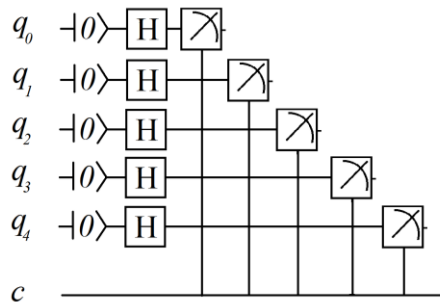
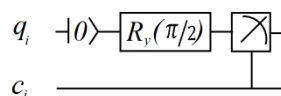


Рисунок 1. Квантовий протокол генерації початкової квантової популяції

Для побудови кожного наступного покоління в процесі еволюції застосовується томографія квантових станів, яка полягає у відновленні амплітуд стану за результатами його багатократного вимірювання. Для ілюстрації розглянемо результат томографії стану кубіта, підготовленого відповідно до наступного квантового протоколу:



Результат томографії квантового стану на квантовому комп'ютері *ibmq-manila* представлено на рис.2.

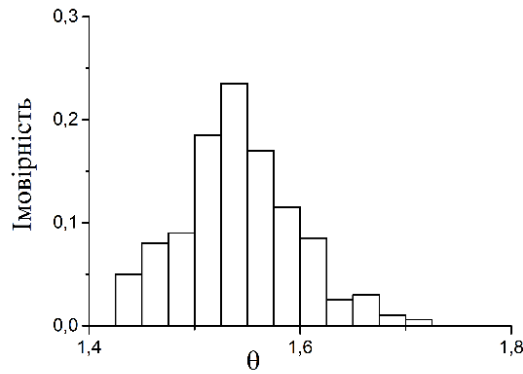
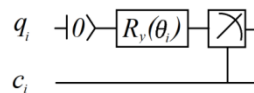


Рисунок 2. Результат томографії стану кубіта при розмірі ансамблю 500 кубітів.

В реалізації еволюційного процесу імовірнісний характер квантової томографії можна розглядати як результат дії оператора квантової мутації, що відповідає за механізм локального пошуку алгоритму. Аналогічно працює «зашумленість» роботи комп'ютера, викликана помилками його роботи в результаті декогерензації квантових станів. Побудова кожної особини в наступному поколінні та її вимірювання реалізовано на квантовому комп'ютері відповідно до наступного протоколу:



Кут θ_i в операторі квантового гейту $R_y(\theta_i)$ визначається результатами квантової томографії попереднього покоління та значенням кута θ_{opt} для найкращої особини популяції відповідно до наступного алгоритму:

$$\theta_i = \begin{cases} \theta_i + \Delta\theta, & \text{при } \theta_i < \theta_{opt} \\ \theta_i - \Delta\theta, & \text{при } \theta_i > \theta_{opt} \end{cases}$$

Величина приросту кута $\Delta\theta$ є параметром роботи алгоритму. Для ефективного глобального пошуку, як показали проведені числові експерименти, його значення повинно, як мінімум, перевищувати $2 \cdot \Delta\theta_{avr}$, отримане в результаті квантової томографії.

Для перевірки ефективності роботи запропонованого підходу з використанням квантового комп'ютера при розмірі популяції 5 кубітів (бо доступними на сьогоднішній день є тільки 5-кубітні квантові комп'ютери) розглянуто задачу функціональної оптимізації ряду стандартних одномірних тестових функцій. Так, наприклад, для функції:

$$f(x) = \sin(x) + \sin\left(\frac{10}{3}x\right), \quad x \in [-2.5, 7]$$

із глобальним максимумом $f(2.296) = 1.728$ динаміку еволюції найкращої особини квантової популяції представлено на рис.3

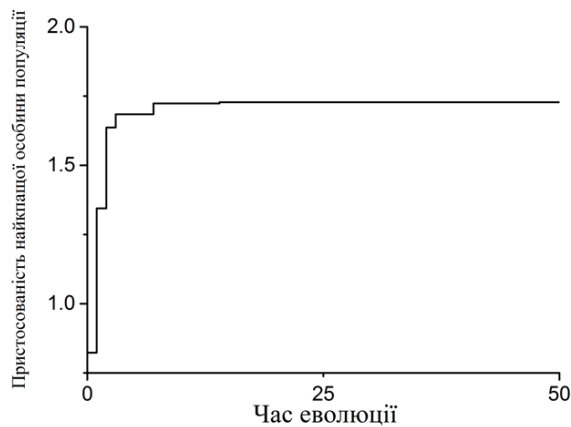


Рисунок 3 Еволюція в часі найкращої особини квантової популяції

V. ОБГОВОРЕННЯ

Результати роботи алгоритму як на доступних квантових комп'ютерах, так і їх симуляторах показують, що не залежно від початкової популяції навіть при 5-ти кубітному її розмірі вона завжди, на протязі не більше 20 циклів еволюції, знаходить глобальний максимум. Оцінка часової ефективності алгоритму провести проблематично, бо генерація кожного нового покоління вимагає очікування в черзі для доступу до квантового процесора.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Перспективним бачиться можливість реалізації квантового клонування без використання томографії для організації еволюційного процесу в часі, однак такий підхід вимагає роботи із квантовими комп'ютерами більшої розмірності, що на сьогоднішній день є достатньо проблематичним.

VII. ВИСНОВКИ

Запропоновано гібридний класично-квантовий алгоритм, у якому класичний комп'ютер виконує функцію підготовки вхідних даних, а всі базові квантові оператори виконуються на квантовому комп'ютері. Для організації еволюційного процесу в часі використовується томографія квантових станів, однак при доступності квантових комп'ютерів більшої розмірності від неї можна буде відмовитись та перейти до використовувати алгоритми наближеного клонування квантових станів. Ефективність запропонованого алгоритму проілюстрована за допомогою набору одномірних тестових функцій для функціональної оптимізації.

ЛІТЕРАТУРА

- [1] A. Narayanan, M. Moore "Quantum-inspired genetic algorithms", 1996, *Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC'96), Nagoya, Japan, 1996*, pp.61–66.
- [2] V. Tkachuk, "Quantum Genetic Algorithm Based on Qutrits and Its Application", *Mathematical Problems in Engineering*, vol. 2018, DOI :10.1155/2018/8614073.
- [3] E. Ballinas, O. Montiel "Hybrid Quantum Genetic Algorithm for the 0-1 Knapsack Problem in the IBM Qiskit Simulator" , *Computación y Sistemas*, Vol. 26, No. 2, 2022, pp. 725–742 doi: 10.13053/CyS-26-2-4253
- [4] IBM Quantum. URL: <https://www.ibm.com/quantum/> (дата звернення 28.03.2022).
- [5] W. Wootters, W. H. Zurek "A Single Quantum Cannot be Cloned", *Nature P. Campbell -Nature Publishing Group*, 1982, vol. 299, p. 802 - 803. - ISSN 0028-0836; 1476-4687-doi:10.1038/299802A0.
- [6] X. Ma, , X. Yuan, , Z. Cao, B. Qi, & Z. Zhang "Quantum random number generation" *Quantum Inf* 2, 16021, doi.org/10.1038/npjqi.2016.21 (2016).

Оцінка якості методів та програмних засобів інтелектуальних адаптивних освітніх веб-систем самонавчання

Марія Дутчак, Микола Пікуляк

*Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація—Завдання оцінки якості методів та програмних засобів інтелектуальних адаптивних освітніх веб-систем самонавчання є багатокритеріальним та багаторівневим, і викликає певні труднощі через слабку формалізованість об'єкта дослідження та відмінності параметрів оптимізації у різних системах. У даній роботі запропоновано уніфікований метод оцінки якості методів та програмних засобів інтелектуальних адаптивних освітніх веб-систем самонавчання, який визначає і структурує найбільш важливі, з точки зору автора, показники якості, та дозволяє оцінити якість даних систем як на рівні окремих показників, так і провести їх зведене оцінювання на різних рівнях ієрархії, враховуючи їх рейтинги.

Ключові слова—*Якість програмного забезпечення, Рейтинг показників якості, Інтелектуальні освітні вебсистеми, Адаптивна траєкторія навчання.*

I. ВСТУП

Інформаційні технології, що застосовуються в інтелектуальних адаптивних освітніх веб-систем самонавчання (ІАОВСС), беруть свій початок в таких галузях як інженерія знань, машинне навчання, інформаційний пошук, Data Mining та в інших галузях штучного інтелекту. ІАОВСС являють собою методологічні та програмно-апаратні рішення, що дозволяють автоматизувати процес пристосування (адаптації) представлення навчального контенту до потреб і параметрів студента [1]. Головним завданням технології адаптації в даних системах є підвищення ефективності процесу навчання. Окрім того, кожна технологія характеризується власними методами, механізмами та технічними засобами, які відповідають різним параметрам оптимізації та способам її реалізації.

При розробці і впровадженні ІАОВСС виникає питання оцінки якості запропонованих рішень в цілому чи їх окремих компонентів. Позитивний ефект від впровадження ІАОВСС залежить як від розроблених проектних рішень, так і від ефективності їх програмної реалізації.

Питання оцінки якості програмного забезпечення ІАОВСС викликає певні труднощі через слабку формалізованість об'єкта дослідження та відмінності параметрів оптимізації у різних системах. Так, наприклад, основну мету автоматизованого навчання можна сформулювати, як переведення студента зі стану, що визначається його вихідною моделлю, в стан, якомога ближчий до цільового стану за оптимальною адаптивною траєкторією навчання (АТН). Очевидним критерієм побудови оптимальної АТН є мінімум відхилень фактичної траєкторії навчання студента від прогнозованої траєкторії і максимальне наближення до цілей навчання. Однак можна запропонувати і багато інших критеріїв оптимальності та якості. Таким чином, завдання визначення рівня оптимальності та якості побудованої АТН є багатокритеріальним.

Також типовою можна вважати ситуацію, коли цілі навчання студента не повністю збігаються з цілями навчання, на які орієнтується ІАОВСС. Тому актуальним завданням є розробка моделей навчання, які враховували б як цілі студента, так і академічні вимоги.

Таким чином, критерії оцінки якості управління навчальним процесом утворюють багаторівневу ієрархічну систему, яка повинна включати в тому числі і формалізацію цілей студентів і академічних вимог. Окрім того, суттєвими особливостями ІАОВСС є значний рівень випадкових зовнішніх збурень на неї, а також високий рівень невизначеності статичних і динамічних характеристик її елементів.

Тому завдання розробки багатокритеріальних та багаторівневих методів оцінки якості ІАОВСС є актуальною і важливою складовою частиною подальшого розвитку і удосконалення даного класу інтелектуальних систем.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

ІАОВСС являють собою мережевий апаратно-програмний комплекс інтелектуального аналізу даних та побудови адаптивного навчального процесу в електронній формі. В англійській літературі близьким до ІАОВСС є термін «Intelligent Online e-Learning Systems». На думку авторів публікації [2], ці системи можна розглядати як інноваційний підхід для надання студенту спеціально-організованої, сфокусованої на студента інтерактивної інформації. Ці твердження розкривають мету ІАОВСС, насамперед, як засобу транспортування індивідуалізованого навчального контенту. Саме тому, на думку деяких вчених, такі системи слід оцінювати з точки зору виконання транспортних функцій – застосовуючи показники якості, властиві телекомунікаційним засобам [3,4].

На думку автора роботи [3] для виконання функцій комп'ютеризованого навчання віддалена система повинна не лише надавати студенту певні сервіси, але й забезпечувати їх якість (Quality of Service, QoS). Засоби, запропоновані науковцем, реалізовано у вигляді інтелектуальної системи, яка використовує апарат нечіткої логіки для прийняття рішень щодо якості телекомунікаційних послуг систем дистанційного навчання за такими параметрами, як пропускна здатність, затримка та втрата пакетів.

На думку автора роботи [5] незважаючи на важливість технічних характеристик мережевої складової ІАОВСС, вказаний підхід не враховує їх призначення як засобу формування та передачі саме індивідуалізованого навчального контенту, дозволяючи моніторити лише можливості окремих технічних засобів щодо передачі потокового відео чи аудіо. Різні за параметрами та можливостями програмне забезпечення ІАОВСС за цими характеристиками буде мати близькі показники. Це стосується і аналізу ефективності ІАОВСС з позиції використання матеріальних-технічних ресурсів.

На думку Кузікова Б.О., ефективність ІАОВСС залежить від того, наскільки вона задовольняє потреби усіх груп користувачів: розробників (авторів та програмістів), споживачів (викладачів та студентів) та адміністративний персонал (адміністраторів, методистів, тощо). В роботі [5] стверджується, що головним ефектом застосування ІАОВСС є зменшення відхилення часу виконання навчальних завдань від планових показників. Для оцінки ефективності за цим параметром було враховано не тільки кількісні, але і якісні показники (складність, тип, відставання від графіку вивчення) за робочою формулою (1.1):

$$k_{plan} = \frac{1}{|C|} \sum_{c \in C} \frac{\sum_{q \in Q_c} \max_{q_i \in q} (p_{q_1} - \max(t_{q_i} - t_q^*, 0) - f_c), 0}{\sum_{q \in Q_c} p_q^*}, \quad (1.1)$$

де C — множина навчальних курсів; Q_c — множина навчальних завдань; дисципліни c , t_q^* та p_q^* — плановий час виконання та максимальний бал для завдання q ; t_{q_i} та p_{q_1} — час та отриманий бал за i -у спробу виконання завдання q ; $f_c = \frac{\sum_{q \in Q_c} p_q^*}{\max(t_q^*) - \min(t_q^*)}$ — штрафний коефіцієнт за затримку на один день при виконанні завдання з дисципліни c . Показник k_{plan} може використовуватись як непряма оцінка знань студентів.

Досить поширеною є практика оцінювання якості ІАОВСС через оцінку рівня отриманих студентом знань. При цьому важливо провести оцінку ІАОВСС саме як засобу організації та транспортування навчального контенту, щоб на неї не впливала якість самого контенту, що ускладнює проведення оцінки [4]. Недоліком безпосередньої оцінки якості знань, як оцінки ефективності ІАОВСС, є неможливість врахування інших параметрів оптимізації та специфічних обмежень системи. Отже, підсумкова оцінка за дисципліну не може слугувати єдиним критерієм ефективності ІАОВСС.

III. МЕТОДИ

На різних етапах впровадження ІАОВСС можуть застосовуватись різні способи оцінки якості ІАОВСС, як специфічного програмного забезпечення. Наприклад, на етапі вибору ІАОВСС можна оцінити її можливості завдяки порівнянню функціональних можливостей різних систем. З цієї точки зору, показовими є методичні матеріали [6], в яких розглянуто 382 критерії для вибору освітніх платформ. При проведенні пробного застосування вдалим може бути застосування спліт-тестування, яке передбачає порівняння показників контрольної та тестових груп, яким були надані для використання продукти, які відрізняються у незначній мірі. Також після пробного застосування ефективність ІАОВСС в цілому чи окремих її модулів можна оцінити провівши опитування ключових груп користувачів. Такий підхід активно використовується в EdX для оцінки якості побудови навчального процесу [7]. Впродовж проходження курсу, користувачам декілька разів пропонується пройти опитування, питання в якому стосуються побудови навчального процесу та самого програмного продукту.

IV. РЕЗУЛЬТАТИ

У склад ІАОВСС входить спеціалізоване програмне забезпечення призначене для збору, зберігання, інтелектуального аналізу та адаптивного виведення даних, тому якість програмного забезпечення даних систем насамперед потрібно розглядати з позиції рівня інтелектуальності та оптимальності побудови навчального процесу.

Оскільки основним завдання ІАОВСС є адаптивне керування навчальним процесом, то при їх оцінці потрібно враховувати показники, за якими здійснюється адаптація навчального процесу, враховуючи, насамперед, мету цього процесу. Мета навчання є основою для формування керуючих впливів на процес навчання. Мета навчання являє собою комплекс цілей студента, викладача та академічних вимог.

Для студентів ефективність від застосування адаптивних інтелектуальних навчальних технологій полягає у оптимізації процесу досягнення поставленої мети, яка може включати такі складові:

- збільшення глибини знань з окремих дисциплін чи тем;
- збільшення обсягів засвоєного навчального матеріалу;
- зменшення часу на засвоєння навчального матеріалу;
- розвиток практичних навичок, тощо.

Саме ці показники, в залежності від мети навчання і параметрів оптимізації мають бути оцінені, бо є прямими показниками ефективності запропонованих рішень.

Провівши аналіз найбільш поширених методів оцінки якості ІАОВСС та враховуючи, що основним завданням ІАОВСС є автоматизація та адаптація процесу оволодіння знаннями з використанням інтелектуального аналізу даних, для оцінки якості методів та програмних засобів даного класу систем був розроблений метод її оцінки, який ґрунтується на показниках якості та їх рейтингах, наведених в табл.1.1.

Таблиця 1.1 – Показники якості методів та програмних засобів ІАОВСС

Характеристика		Підхарактеристика	Рейтинг
Назва	Шифр	Назва	
Функціональність	Ф	релевантність побудованої адаптивної траєкторії навчання	В
		цілісність представлення адаптованого навчального матеріалу	В
		адекватність моделі прогнозування навчальних досягнень	В
		якість контролю знань	В
		вплив на рівень засвоєння навчального матеріалу	В
		оцінка досягнення студентами поставленої мети	В
Зручність користування	З	зручність та зрозумілість інтерфейсу користувача	С
		автоматизація процесу формування бази знань	С
Ефективність	Е	ефективність використання праці викладачів та експертів	В
		ефективність використання технічних засобів	С
		час відповіді системи	С
Надійність	Н	відновлюваність даних	В
		цілісність бази знань	В
Переносимість	П	легкість розгортання	Н
		кросплатформність	Н

Позначення у табл. 1.1 наступні: В (високий), С (середній) та Н (низький) – рейтинги показників якості програмних продуктів; значення показників якості є уніфікованими та нормованими шляхом переведення їх в діапазон від 0 до 1.

Якість методів та програмних засобів ІАОВСС можна оцінювати як числовим значенням окремих вихідних показників (критеріїв, характеристик, метрик), так і зведеною оцінкою всієї сукупності показників із врахуванням рейтингу їх важливості. Оцінювання проводиться безпосередньо експертами за частковими критеріями методами експерименту, опитування, оцінювання окремих показників роботи системи тощо, які в подальшому об'єднуються з іншими частковими оцінками цього рівня ієрархії і таким чином утворюються узагальнені (проміжні) оцінки на вищому рівні ієрархії. Потім відбувається об'єднання на вищому рівні і т. д. Такий процес продовжується до отримання потрібної кінцевої оцінки якості.

V. ВИСНОВКИ

На основі проведеного дослідження можна зробити висновки, що завдання оцінки якості методів та програмних засобів ІАОВСС є багатокритеріальним та багаторівневим, і викликає певні труднощі через слабку формалізованість об'єкта дослідження та відмінності параметрів оптимізації у різних системах.

У даній роботі запропоновано уніфікований метод оцінки якості методів та програмних засобів ІАОВСС, який визначає і структурує найбільш важливі, з точки зору автора, показники якості, та дозволяє оцінити якість даних систем як на рівні окремих показників, так і провести їх зведене оцінювання на різних рівнях ієрархії, враховуючи їх рейтинги.

Отже, розробка нових і ефективних методів оцінки якості методів і програмних засобів ІАОВСС є актуальною і важливою складовою частиною подальшого розвитку і удосконалення даного класу інтелектуальних систем, оскільки дає можливість проаналізувати і обрати найбільш оптимальні та перспективні технології їх розвитку.

ЛІТЕРАТУРА

1. М. С. Дутчак, “Методи та програмні засоби автоматизованої побудови адаптивної траєкторії навчання,” *Вісник Вінницького політехнічного інституту*, № 2. (149). сс.58-66, 2020, doi: <https://doi.org/10.31649/1997-9266-2020-149-2-58-66>.
2. J.R. Corbeil, B.H. Khan and M.E. Corbeil, “Microlearning in the Digital Age,” in *Routledge*, New York, NY, USA, 2021, p. 68.
3. Г.Г. Маклакова, “Интеллектуальное управление в распределенной системе дистанционного обучения,” *Управляющие системы и машины*, № 2. сс. 73-76, 2009. Режим доступа: http://nbuv.gov.ua/UJRN/USM_2009_2_12
4. А. С. Довбиш, А. В. Васильев і В. О. Любчак, “Интеллектуальні інформаційні технології в електронному навчанні: монографія,” *СумДУ*, 2013, с. 177.
5. Б. О. Кузіков, “Підходи до оцінки ефективності адаптивної системи дистанційного навчання,” *Вісник Сумського державного університету. Сер.: Технічні науки*, № 3. сс. 67-74, 2013. Режим доступа: http://nbuv.gov.ua/UJRN/VSU_tekh_2013_3_11
6. B. Brandon and C. Holcombe, “Tips for Selection of an LMS or LCMS,” in *The eLearning Guild*, December 2006, p. 58.
7. Y. Chaabi and Y. Al-Ashmoery, "Development of a Learning Analytics extension in Open edX," in *2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI)*, 2021, pp. 1-6, doi: 10.1109/MTICTI53925.2021.9664754

Програмні аспекти виявлення академічного плагіату

Микола Пікуляк, Марія Дутчак

*Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація— Виконано порівняльний аналіз методів, що застосовуються у відомих алгоритмах перевірки текстів на предмет виявлення запозичень та обчислення відсотку унікальності тексту. Виділено переваги та недоліки їх використання для перевірки наукових робіт із різних предметних областей. Проведено дослідження на основі методу статистичного групування результатів перевірки кваліфікаційних робіт у Прикарпатському національному університеті протягом двох навчальних років з використанням програмних сервісів Unicheck та Plagiat.pl, що дозволило виявити найбільш типові способи порушення академічної доброчесності студентами та визначити подальші напрямки щодо розробки вдосконалених методів та алгоритмів автоматизованого аналізу тексту. Визначено перспективи подальших досліджень, направлених на розробку сучасних програмних рішень з метою пошуку академічного плагіату.

Ключові слова— *Академічна доброчесність, Антиплагіатні програми, Методи аналізу тексту, Алгоритми створення антиплагіатних сервісів, Статистична методологія.*

I. ВСТУП

Впровадження принципів академічної доброчесності в навчальному процесі та науковій діяльності є важливою умовою успішного розвитку вищої освіти в Україні та запорукою визнання наукових досягнень українців світовою спільнотою. З однієї сторони зрозуміло, що всі нові наукові результати повинні базуватись та відштовхуватись від вже відомих. Оскільки для того, щоб створити щось нове в науці, необхідний детальний аналіз вже досягнутих успіхів у тій чи іншій предметній області. З іншої сторони, потрібно унеможливити використання у нових дослідженнях академічного плагіату, під яким розуміють оприлюднення наукових (творчих) результатів, отриманих іншими особами, як результатів власного дослідження без зазначення авторства [1].

Особливо поширеною тенденцією маніпулювання чужими ідеями та результатами в останній час є використання «рерайту» (тобто грамотного переписування оригінальної роботи) та «копіпасти» (копіювання фрагментів із декількох джерел). Якісне використання користувачами таких інструментів забезпечує внесення структурних відмінностей в текстову частину документу, хоча сам зміст та наукова складова залишаються ті ж самі, що і в джерелі.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

З метою виявлення та запобігання академічним правопорушенням на сьогодні розроблено чимало платних та безкоштовних сервісів (антиплагіатних програм), які дозволяють автоматично відстежувати текстові запозичення та визначати співпадіння й відсоток унікальності тексту в мережі Інтернет і Репозитарії. В основі їх розробки, як правило, лежать власні алгоритми аналізу текстів та певні методи пошуку прихованого плагіату, найбільш поширеними з яких є «синтаксичні» та «лексичні» методи.

До «синтаксичних» відносяться деякі прості методи синтаксичного аналізу документів, метод шинглів (алгоритм А. Бродера), Long Sent та Heavy Sent [2]. Такі методи використовують сучасні криптографічні протоколи, математичні обрахунки у яких виконують відповідні функції хешування. При цьому під хешуванням розуміють

відображення, що перетворює повідомлення довільної довжини у хеш-код фіксованої довжини n , тобто:

$$F : \{0,1\}^* \rightarrow \{0,1\}^n$$

Зокрема, у нескладних синтаксичних методах, як правило, обчислюють хеш-функцію від конкатенації двох найдовших речень, знайдених у тексті та обраховують відношення числа входжень певного слова до усієї кількості слів у тексті. Такі підходи є не точними та можуть бути використані тільки для попередньої оцінки.

У методі «шинглів» використовують послідовності фіксованої довжини, які складаються з сусідніх слів («шингли»). Від кожної послідовності («шингла») береться хеш-сума і створюється набір хеш-сум. Відсоток схожості у двох документах буде тим більшим, чим частіше перетинаються їхні множини шинглів. У кожній множині обчислюються дактилограми (тобто текстові підрядки фіксованої довжини) за допомогою взаємно-однозначних і незалежних функцій, що використовують випадкові набори простих поліномів. Головним недоліком алгоритму шинглів є зменшення продуктивності під час перевірки текстів великих об'ємів, оскільки збільшення кількості шинглів призводить до зростання операцій для виконання відповідних порівнянь.

У зв'язку з тим, що користувачі постійно знаходять нові способи обходу дії антиплагіатних програм, тому алгоритми пошуку репрезентативних підмножин постійно вдосконалюються. Сьогодні алгоритм А. Бродера отримав такі вдосконалення:

- алгоритм Карпа-Рабіна – вибирає тільки ті шингли, кількість «дактилограм» яких ділиться без залишку на деяке число m . Це дозволяє збільшити продуктивність при обробці великих обсягів тексту. Основний недолік – залежність вибірки від довжини документа;

- метод Long Sent, у якому документ спочатку розбивається на речення, які упорядковуються за зменшенням довжини, представленої кількістю слів (при рівності довжин – у алфавітному порядку). Потім вибираються та зчіплюються у рядок в алфавітному порядку два найдовші речення. В якості сигнатури документу обчислюється контрольна сума отриманого рядка. Головний недолік методу – невисока швидкість та низький процент плагіату при перевірці великого тексту;

- метод Heavy Sent – ідея методу схожа до методу Long Sent: документ ділиться на речення і два з них зчіплюють в рядок в алфавітному порядку. Однак вибір здійснюється дещо інакше: враховують вагу кожної пропозиції, що являє собою суму ваг слів, які входять до нього, і вибір припадає на дві «найважчі» пропозиції. Контрольна сума, отримана після зчеплення рядків приймається за сигнатуру документа;

- метод Moodle Crot [3] – спочатку із тексту видаляються слова до трьох символів, а потім всі небуквені знаки (дефіси, крапки, пробіли і т.д.). Потім отриманий у такий спосіб суцільний ланцюжок букв з певним кроком n «нарізається» на частини по N символів у кожній та від кожної частини береться хеш-функція, результат якої використовується для аналізу.

Прикладами лексичних методів є метод TF-IDF та Метод Коудури [4]. Основна ідея використання таких методів полягає у використанні для визначення унікальності документу спеціальних словників.

Зокрема у методі TF-IDF складається словник, що ставить кожному слову число документів, в яких воно зустрічається. Потім будується частотний словник документа і для кожного слова обчислюється його «вага». Недоліки методу – в сигнатуру не включаються рідкісні слова та низький процент плагіату при перевірці великого тексту.

У методі Коудури лексичний словник будується на основі дактилограм I-Matcha [5], за числовими значеннями яких обраховують унікальність відповідних документів. Перевагою цього методу є більш висока обчислювальна ефективність ніж в алгоритмі А. Бродера. Недоліком є нестійкість до незначних змін змісту документу. Особливо корисним його застосування є під час вебпошуку та фільтрації спаму.

Серед сучасних методів машинного навчання слід виділити метод з використанням граф-трансформерів [6], перевагою якого є можливість за рахунок так званого механізму уваги виконувати обробку будь-якого положення слова в послідовності входу та визначати міжмовні текстові збіги. Механізм уваги дозволяє у моделі декодування використовувати вектор стану кожного вхідного англійського слова, а не лише крайнього (як це робиться в класичній моделі рекурентних нейронних мереж). Це забезпечує більш ефективний спосіб пошуку значення українського слова в кожному з векторів стану англійського входу.

Також сьогодні ведуться спроби розробки програм штучного інтелекту, які на основі методів нейронних мереж та генетичних алгоритмів дозволяють проводити аналіз документів не тільки на предмет порівняння текстової частини, але й визначають плагіат зображень, моделей, конструкцій та програмних кодів.

III. МЕТОДИ

В роботі використана статистична методологія аналізу результатів дослідження, яка дозволяє на основі математичної обробки отриманих результатів оцінити динаміку зміни якості кваліфікаційних робіт, визначити найбільш типові академічні правопорушення студентами та провести статистичне узагальнення щодо розробки вдосконалених методів та алгоритмів створення антиплагіатних програм.

Перевагою застосування даної методики є можливість обчислити узагальнюючі показники зміни діапазону коефіцієнтів подібності, що описують дані окремих навчальних підрозділів та здійснити аналіз і прогнозування використаних алгоритмів аналізу текстових збігів.

Зокрема в дослідженні застосовується метод статистичного групування [7], який дає змогу виявити об'єктивний стан предметної області, простежити перехід кількісних показників у якісний аналіз текстової інформації, отримати дані про зв'язки між досліджуваними показниками, їх співвідношення в загальній сукупності та оцінити закономірності їхнього розвитку.

Методологічна сутність запропонованого групування полягає у об'єднанні окремих діапазонів коефіцієнтів подібності у групи за числовими значенням, що дозволяє здобути інформацію про розміри окремих груп та прослідкувати залежність між ознаками, що покладені в основу групування.

IV. РЕЗУЛЬТАТИ

В основу дослідження покладено два програмні продукти Unicheck та Plagiat.pl, які використовуються у Прикарпатському національному університеті для перевірки на плагіат студентських та наукових робіт та з виробниками яких укладено відповідні угоди Міністерством освіти і науки України (МОН).

Результатом роботи даних сервісів є формування підсумкового звіту (рис. 1):

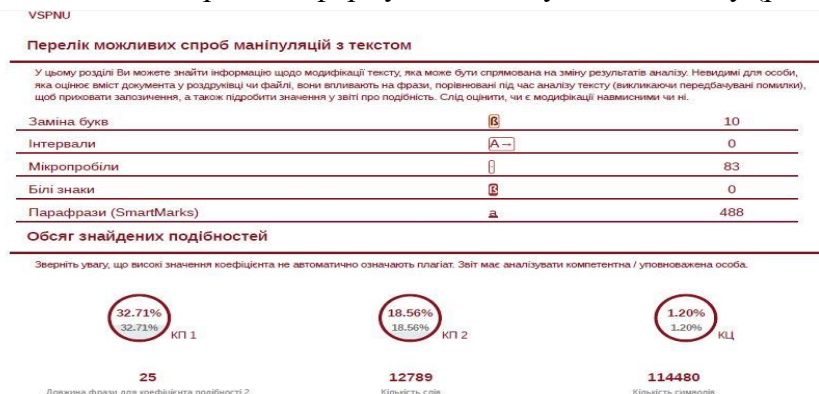


Рисунок 1 – Вікно результатів обробки тексту на плагіат

Основними елементами звіту служать два коефіцієнти подібності КП1 та КП2, які дають можливість визначити наступні типи запозичень:

- процентну частку аналізованого документу, що є ідентичною виявленим джерелам;
- список джерел фрагментів, позначених як скопійовані;
- повний текст перевіреного документу, з чітко визначеними виявленими запозиченнями;
- тривожний сигнал, що інформує про викривлення в тексті, в результаті використання різних алфавітів, що може вказувати на спробу приховання плагіату.

Беручи до уваги звіти, сформовані за результатами перевірки кваліфікаційних робіт у програмних сервісах Unicheck та Plagiat.pl було виконано аналіз дипломних та магістерських робіт протягом двох навчальних років (2020-2021 та 2021-2022 н. р.) у Прикарпатському національному університеті. В результаті проведеного дослідження були отримані статистичні дані, які відображають динаміку зміни якості навчальних матеріалів та відсоток текстових запозичень (рис. 2):

Динаміка						
Підрозділ	2020-2021		2021-2022		Прогрес	
	% якості	Середній %	% якості	Середній %	% якості	Середній %
Факультет математики та інформатики	74,51	19,25	91	10,7	16,49	-8,55
Факультет історії, політології і міжнародних відносин	80,65	17,66	84,2	12,6	3,55	-5,06
Фізико-технічний факультет	81,08	15,3	83,8	13,6	2,72	-1,7
Факультет фізичного виховання і спорту	90,62	17,29	93,3	13,6	2,68	-3,69
Факультет природничих наук	49,07	34,66	84,7	14,9	35,63	-19,76
Навчально-науковий Інститут мистецтв	58,97	23,01	80,4	15,2	21,43	-9,81
Факультет філології	75,86	20,05	80	18,4	4,14	-1,65
Економічний факультет	71,62	23,2	75,6	19,3	3,98	-3,9
Факультет психології	100	0	76,7	19,8	-23,3	19,8
Факультет туризму	37,5	38,6	72,9	20,5	35,4	-18,1
Педагогічний факультет	63,46	22,86	69,9	25,2	6,44	2,34
Інститут післядипломної освіти та довузівської підготовки	67,74	25,75	57,4	31,4	-10,34	5,65

Рисунок 2 – Динаміка зміни якості кваліфікаційних робіт у програмних сервісах Unicheck та Plagiat.pl

На основі застосування методу статистичного групування отримано динаміку зміни діапазону коефіцієнтів подібності за два роки (рис.3):

Динаміка (літні сесії, ОР бакалавр)			
Діапазон коефіцієнтів подібності		2020-2021	2021-2022
1	0% - 20%	50,6%	61,0%
2	20% - 40%	24,4%	19,4%
3	40% - 60%	13,0%	9,4%
4	60% - 100%	11,9%	10,2%

Рисунок 3 – Динаміка зміни діапазону коефіцієнтів подібності

Отримані статистичні обчислення дозволили з однієї сторони зробити висновок про ефективність та подальшу доцільність застосування сервісів Unicheck та Plagiat.pl для перевірки кваліфікаційних робіт. З другої сторони на основі застосування даного методу

було виявлені взаємозв'язки і залежності між отриманими статистичними обрахунками та визначені типові способи використання плагіату студентами.

Зокрема в результаті проведеного детального аналізу досліджуваних робіт на факультеті математики та інформатики (опрацьовано звіти подібності чотирьох академічних груп (74 студенти)) було виявлено наступні, найбільш типові способи використання рерайтингу та копіпасти студентами:

- дослівне копіювання (зі списуванням включно);
- синонімічні заміни (синонімайзинг) із термінологічними замінами включно;
- зміна порядку слів;
- переклад з іншої мови;
- зміна ключових слів і фраз при збереженні основного змісту;
- введення сполучних слів і словосполучень, перехідних фраз, невеликих коментарів;
- плагіат даних (результати експериментів, емпіричних досліджень, розрахунків);
- плагіат зображень (фото, діаграми, схеми, креслення та ін.);
- запозичення програмних кодів.

З метою виявлення вказаних вище запозичень та унеможливлення подальшого використання вважаємо за необхідне використання вдосконалених методів та алгоритмів створення антиплагіатних програм за рахунок:

- перед обчисленням хеш-суми у методі шинглів виконувати сортування слів у шинглі, що дозволить покращити пошук запозичень у випадках, коли слова у тексті проаналізованої роботи були переставлені місцями;

- для збільшення ефективності використовувати модифіковані версії алгоритму шинглів, а саме:

- алгоритм MinHash [8], який кожній хеш-функції ставить у відповідність число – найменше значення для всіх шинглів досліджуваного рядка. В результаті для обробки отримуємо одновимірний масив (вектор) хеш-кодів довжин (короткі числові образи документу), завдяки чому значно зростає швидкість пошуку ймовірних дублікатів;
- алгоритм SimHash [9] – текст розбивається і кожен фрагмент, представлений у вигляді бінарного вектора хешується вибраною функцією. За рахунок того, що замість перевірки всього набору виконується порівняння окремих сигнатур, значно зростає швидкість алгоритму;

- щоб визначити заміну літер з кирилиці на латиницю на етапі нормалізації тексту застосовувати так званий метод синонімізації слів [3], який передбачає використання спеціальних словників (хоча проблемою використання цього методу залишається на сьогодні досить низька якість словників з української мови, які знаходяться у відкритому доступі, а якість комерційних словників важко перевірити через обмежений до них доступ);

- використовувати перехресну перевірку з допуском збігів, згідно з якою у «шингли» поміщати не хешсуми, а безпосередньо самі ланцюжки слів; а для більш детальної перевірки додатково використовувати словники синонімів та антонімів;

- коли потрібна більша точність перевірки застосовувати локальні алгоритми аналізу текстів як додаткові до методу шинглів (наприклад, Moodle Crot);

- для покращення роботи усіх методів на етапі нормалізації потрібно видаляти відомі словосполучення, враховувати синоніми, приводити літери до одного регістру та алфавіту;

- для пошуку міжмовних текстових збігів використовувати сучасні підходи з використанням моделей-трансформерів та граф-трансформерів;

- для запобігання копіювання чужих програмних кодів рекомендувати перевіряти програмні продукти з використання інструменту Copyleaks Programming Code Checker, який на основі штучного інтелекту може знаходити схожі шаблони коду в Інтернеті та базах даних.

V. ОБГОВОРЕННЯ

Як показує проведений аналіз сьогодні для перевірки наукових та навчальних праць наявність академічного плагіату доступно цілий ряд сервісів. Використання таких систем є вимогою часу та МОН. Проте дуже важливо розуміти, що не існує програм, які забезпечують 100-відсоткову перевірку тексту на плагіат. Антиплагіатні рішення не виявляють академічного плагіату, а лише допомагають його шукати.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Подальшим напрямком дослідження буде розробка програмного сервісу на основі застосування нових алгоритмів автоматизованого аналізу, який дозволить більш якісно проводити перевірку студентських робіт та наукових праць викладачів для виявлення академічного плагіату.

VII. ВИСНОВКИ

В роботі виконано аналіз відомих методів та алгоритмів, які використовуються для програмного пошуку текстових запозичень.

На основі проведеного статистичного дослідження щодо аналізу студентських робіт на предмет виявлення академічного плагіату визначено динаміку зміни якості кваліфікаційних робіт та виділено недоліки в існуючих антиплагіатних програмах, що дозволило стверджувати про доцільність подальших розробок сучасних програмних рішень та проведення еволюційно-проектних вдосконалень в області дотримання академічної доброчесності.

ЛІТЕРАТУРА

[1] Верховна Рада України. Закон України про вищу освіту: документ 1556-VII (28 груд. 2014, зі змінами 25 верес. 2020). [Електронний ресурс]. Режим доступу: <https://zakon.rada.gov.ua/laws/show/2145-19#Text>. Дата звернення: Жовт. 26, 2022.

[2] О.А. Куценко, О.А. Сердюк, "Огляд алгоритмів пошуку плагіату у програмному коді," Вісник Черкаського національного університету імені Богдана Хмельницького, Вип. 1, С. 75 – 85, 2019. DOI 10.31651/2076-5886-2019-1-75-85.

[3] А.О. Білощицький, О.В. Діхтяренко, "Ефективність методів пошуку збігів у текстах," Управління розвитком складних систем, № 14, С. 144–147, 2013.

[4] A. Kolcz, A. Chowdhury, J. Alspector, "Improved Ro-bustness of Signature-Based Near-Replica Detection via Lexi-con Randomization, " *KDD 2004*, 22-25 August, 2004, Seattle, Washington, USA.

[5] A. Chowdhury, O. Frieder, D. Grossman, M. McCabe, "Collection statistics for fast duplicate document detection, " *ACM Transactions on Information Systems (TOIS)*, vol. 20, Issue 2, April 2002.

[6] The Illustrated Transformer. [Електронний ресурс]. Режим доступу: <https://jalammar.github.io/illustrated-transformer/>. Дата звернення: Лист. 08, 2022.

[7] А.Т. Мармоза. Теорія статистики: Підручник. 2- ге вид. перер. та доп, К.: Центр учбової літератури, 2013, 592 с.

[8] A. Shrivastava, "Simple and efficient weighted minwise hashing, " *Advances in Neural Information Processing Systems*, pp. 1498–1506, 2016.

[9] Moses S. Charikar, "Similarity estimation techniques from rounding algorithms," *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, 2002, p. 380.

Machine Learning Based Animal Emotion Classification Using Audio Signals

Mariia Slobodian

Wallee AG, Winterthur, Switzerland

Mykola Kozlenko

SoftServe, Austin, USA

Abstract—This paper presents the machine learning approach to the automated classification of a dog's emotional state based on the processing and recognition of audio signals. It offers helpful information for improving human-machine interfaces and developing more precise tools for classifying emotions from acoustic data. The presented model demonstrates an overall accuracy value above 70% for audio signals recorded for one dog.

Keywords—*acoustic features, audio signals, dog vocalization analysis, machine learning, deep learning, artificial neural network, mobile application, cepstral coefficients, sound segmentation.*

I. Introduction

Scientists suggest that canines are far more intelligent than people realize. Over the years, there have been a lot of publications about studies focused on dogs. Studies have revealed that dogs demonstrate advanced emotional intelligence. They are equipped to perform great deeds for their owners and have a high emotional sensitivity. Accordingly, it is reasonable to assume that humans have long felt the desire to speak with dogs. In addition, there have been a lot of attempts to handle and process the audio signals. It's interesting how vocal emotion recognition software functions because it resembles facial expression detection software in many ways. Applying human voice emotion recognition techniques to dog recordings also shows promising results. The best accuracy so far is ultimately insufficient for use in practical applications.

II. Related works

A number of similar research works in the field of modern human-machine interfaces have used recognition systems, such as human voice recognition by emotion and gender, and they have shown strong accuracy results. Emotional prosody can be used to identify emotions, according to research on human speech. [1]. This study used comparisons with other live species to examine the evolution of speech from a comparative and phylogenetic perspective. People can recognize barks that are related to circumstances such as being willing to pay, being angry, or being aggressive [2]. It was proved that dog barks are the most well-known animal voice to humans [3]. Furthermore, barks have specific acoustic features [4]. Changes in an animal's emotional state should have an impact on the muscle systems that govern its vocal apparatus, changing the acoustic characteristics of vocal transmission, especially when the animal engages in safety defense [4].

A preliminary investigation of machine learning has revealed that auditory descriptors are appropriate for detecting bark context, with a classification efficiency of 43% [5]. Different acoustic features that are known to reflect human emotions were compared in order to divide canine bark patterns according to their context or observed emotion. Publications [6], [7], [8] offer a method that effectively separates relevant information from useful signals and random noise using deep learning and artificial neural networks. Applying those methods to categorize the emotions of dogs seems promising.

III. Methodology

During our research, we identified two key methods with typical data preprocessing approaches such as Mel-frequency cepstral coefficients (MFCC) and Extractor Discovery System (EDS). It demonstrates that while MFCC is primarily used to obtain information from audio files and depict it using pictures, EDS iteratively generates descriptor generations and can mark the best-performing descriptors from one generation as seeds for subsequent ones. To solve this problem, EDS employs a range of artificial intelligence tools. First off, a type system forbids the creation of descriptors that include mistakes. In order to further elaborate, new descriptors are employed as the second phase. On the other hand, by using some effective classifiers, MFCC gives us 13 relevant coefficients at the end that we can use with appropriate machine learning techniques. But in this research, we do not use any preprocessing and feature engineering steps. Audio signals pass directly to the convolutional neural network (CNN).

We found that the ideal set of emotions contains five types of feelings (Fig. 1), which are fewer compared to human measurements. Those emotion classes served well our needs in model classification. Furthermore, it requires extra effort to analyze the dynamic movement of the vocal folds. The following voice cues could be used to identify the vocal characteristics: timbre, amplitude, loudness, intensity, and spectral composition.

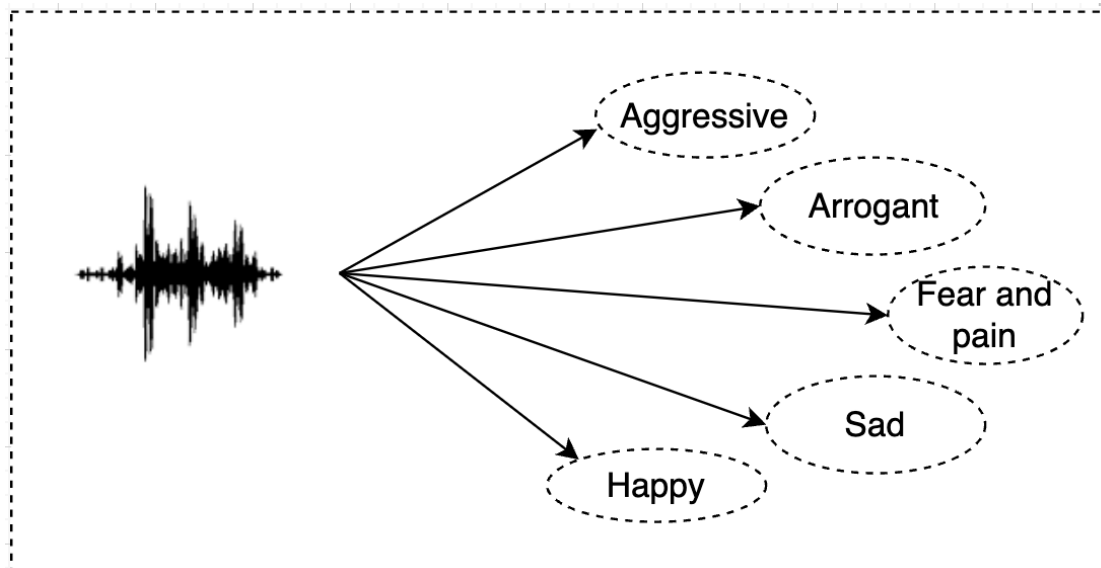


Fig.1. Emotion classes

The structure of the developed deep learning model is shown in Fig. 2. It gets the fragment of the audio signal at the input. The length of the signal is a hyperparameter (12 000 samples, for example). It contains two hidden convolutional layers with a ReLU activation function and one fully-connected layer with five output neurons and softmax activation. Batch normalization is used between all layers. The model outputs five confidence values for each emotion class (Fig. 2.). The model was trained with 4000 signal fragments. The size of the validation set is 800 records.

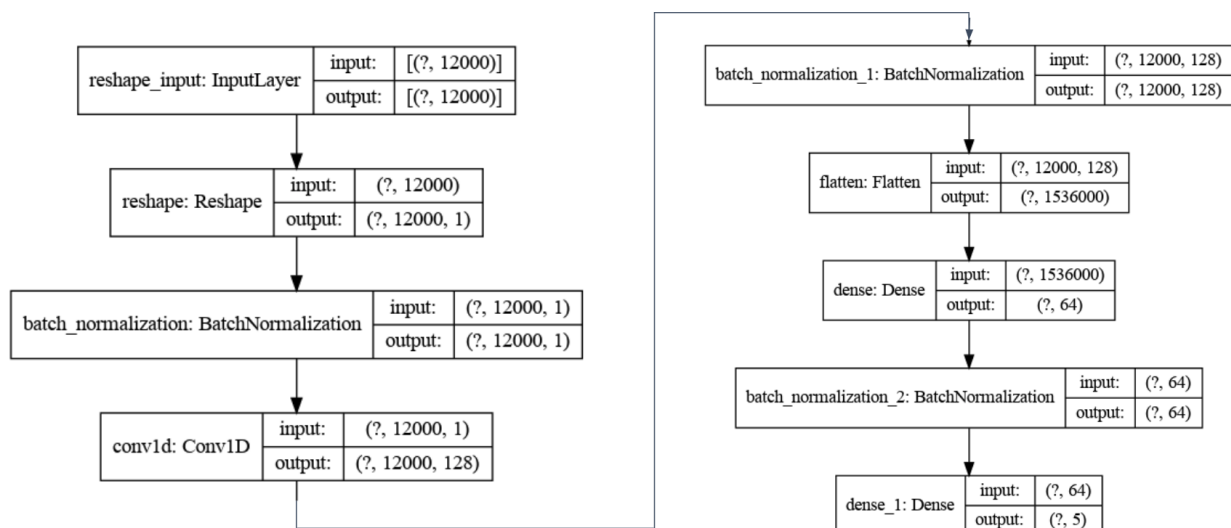


Fig. 2. Structure of the deep learning model

The model performance assessment is performed as a post-predict evaluation. It demonstrates how closely the predictions represent the actual ground truth. We use a widely-used classification report that outputs class-wise precision, recall, f1-score as well as overall accuracy. The size of the test set is 1500 records.

IV. Results

The classification report is shown in Fig.3. The overall accuracy is about 72% (model trained from one dog).

	precision	recall	f1-score	support
aggressive	0.73	0.81	0.77	280
arrogant	0.70	0.61	0.65	288
fear_and_pain	0.66	0.93	0.77	302
happy	0.80	0.54	0.64	329
sad	0.76	0.74	0.75	301
accuracy			0.72	1500
macro avg	0.73	0.73	0.72	1500
weighted avg	0.73	0.72	0.72	1500

Fig. 3. Classification report

V. Discussion

The variety of feelings could be recognized by us, but in comparison to positive ones, the proportion of negative vocalizations was dominant. It can be challenging to elicit neutral vocalizations because anger and playfulness are typically active emotions. Despite that, the joyful class of barking is still the most recognizable despite the limited data. It is challenging to train models because of the lack of labeled data of good quality. Different outcomes were seen when the same model was applied to the entire group of dogs as opposed to each individual dog.

VI. Future research

The experiments have demonstrated the potential of using machine learning models to distinguish between various bark properties. However, it seemed that the attained accuracy wasn't good enough for applications in the real world. Additional training and improvement are needed to achieve greater results. Soon, we'll investigate how well-known feature extraction techniques can be linked with deep learning strategies as well as the direct usage of raw signals as features. It could greatly reduce the number of computer resources needed and streamline the data processing pipeline. Analysis of the potential use of recurrent neural networks and architectures is also necessary.

VII. Conclusion

We examined the possibility of distinguishing dog bark using acoustic feature sets and affective computing based on context, perceived emotion, and intensity. Moreover, we offered recommendations for how to enhance the examined techniques for classifying emotions. It must be emphasized that the approach analysis offered here could make a substantial impact and excite business interest.

VIII. Acknowledgment

The authors sincerely acknowledge the assistance of the experts of the Department of Information Technology of the Vasyl Stefanyk Precarpathian National University for their technical support during the real analysis as well as their professional guidance for consultations.

IX. Disclosures

The authors declare that there are no conflicts of interest related to this paper.

References

- [1] W.T. Fitch, "The biology and evolution of speech: A comparative analysis," *Annual Review of Linguistics*, vol. 4, pp. 255–279, 2018, doi: 10.1146/ANNUREV-LINGUISTICS-011817-045748.
- [2] S. Hantke, N. Cummins, and B. Schuller, "What is my Dog Trying to Tell Me? the Automatic Recognition of the Context and Perceived Emotion of Dog Barks," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 5134–5138, doi: 10.1109/ICASSP.2018.8461757.
- [3] D. U. Feddersen-Petersen, "Vocalization of European wolves (*canis lupus lupus* L.) and various dog breeds (*canis lupus* F. Fam.)," *Archives Animal Breeding*, vol. 43, no. 4, pp. 387–398, 2000, doi: 10.5194/AAB-43-387-2000.
- [4] S. Yin, "A new perspective on barking in dogs (*canis familiaris*.)" *Journal of Comparative Psychology*, vol. 116, no. 2, pp. 189–193, 2002, doi: 10.1037/0735-7036.116.2.189.
- [5] C. Molnar, F. Kaplan, P. Roy, F. Pachet, P. Pongracz, A. Doka, and A. Miklosi, "Classification of Dog Barks: A Machine Learning Approach," *Animal Cognition*, vol. 11, pp. 389–400, 2008, doi: 10.1007/S10071-007-0129-9
- [6] M. Kozlenko, I. Lazarovych, V. Tkachuk, and V. Vialkova, "Software Demodulation of Weak Radio Signals using Convolutional Neural Network," *2020 IEEE 7th International Conference on Energy Smart Systems (ESS)*, 2020, pp. 339–342, doi: 10.1109/ESS50319.2020.9160035

- [7] M. Kozlenko and V. Vialkova, "Software Defined Demodulation of Multiple Frequency Shift Keying with Dense Neural Network for Weak Signal Communications," *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2020, pp. 590-595, doi: 10.1109/TCSET49122.2020.235501
- [8] I. Lazarovych et al., "Software Implemented Enhanced Efficiency BPSK Demodulator Based on Perceptron Model with Randomization," *2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2021, pp. 221-225, doi: 10.1109/UKRCON53503.2021.9575458.

The third initial-boundary value problem for the nonlinear equation of heat conduction with deviation of the argument

Yaroslav Drin'¹, Iryna Drin'², Roman Drin'³

^{1,3} *Yuriy Fedkovych Chernivtsi National University*

² *Chernivtsi Trade and Economic Institute of the State Trade University*
Chernivtsi, Ukraine

Abstract—The third initial-boundary problem for the heat conduction equation with the inversion of the argument are considered. The Green's function of considered problem are determined. The theorem about the Poisson integral limitation is proved. The theorem declared that the Poisson integral determines the solution of the third initial-boundary problem considered and proved.

Keywords—Heat Nonlinear Equation, Boundary Value Problem, Green Function, Deviation Argument.

I. INTRODUCTION

In this paper the solution of the third initial-boundary value problem for the heat conduction equation with the inversion of the argument is found. The process of heat conduction was first described by Jean Baptiste Joseph Fourier (1768 – 1830) in 1807 in the work "Equations with partial derivatives for heat conduction in solids". A description of the results of other scientists who studied and developed this theory is presented in [1, 2]. Thermal conductivity is the molecular distribution of thermal energy in various solids, liquids and gases due to the difference in temperature and due to the fact that the particles are in direct contact with each other. Based on different criteria, models of heat conduction processes are divided into two groups of models using integral and fractional order derivatives. Models with derivatives of an integer order are stationary and non-stationary.

Non-stationary models of one-dimensional heat conduction are described by the equation of heat conduction [3]. Different methods of solving this problem are described in [2]. Applied aspects of such problems are described in [3, 4, 5].

If the evolution of the concentration of impurities, point defects, and the temperature field is studied, then the corresponding transfer coefficients are not constant values. Processes with spatially dependent transfer coefficients or the desired field are well studied [6, 7].

Our work examines the third initial-boundary value problem for the non-homogeneous nonlinear heat conduction equation with argument deviation. The first and second initial-boundary value problems are studied in [8, 9].

II. FORMULATION OF THE THIRD INITIAL-BOUNDARY VALUE PROBLEM

Let $a > 0, b > 0$ be real numbers; $x \in R^+, t \in R^+$ are independent variables; f, φ, μ are known continuous functions; $u(x, t)$ is the desired function that describes the evolution of the system defined on the semi-axis $x \in R^+$ for all $t \in R^+$. We will study the problem

$$u_t = a^2 u_{xx} + f(x, t, u(x, t-h)), x > 0, t > h, \quad (1)$$

$$u(x, t)|_{0 \leq t \leq h} = \varphi(x, t), x \geq 0, \quad (2)$$

$$u_x(0, t) - \gamma u(0, t) = \mu(t), t \geq h, \gamma > 0, \quad (3)$$

which is the third initial-boundary value problem, where the functions $\varphi(x, t) \in C(R^+ \times \{0 \leq t \leq h\})$ is initial function, $\mu(t) \in C(R_n^+)$ is boundary function, $R_n^+ \equiv \{t; t \geq h\}$, $R^+ \equiv \{x; x \geq 0\}$, $f(x, t, u) \in$

$\in C(R^+ \times R_h^+ \times R)$ is the inhomogeneity of equation (1) is well known. If a smooth solution of the problem (1) – (3) is sought up to the limit, then the initial and boundary functions must be consistent $\varphi_x(x, h) - \gamma\varphi(x, h) = \mu(h)$.

III. THE STEPS METHOD

Let $h \leq t \leq 2h$, then $t - h \geq 0$, $u(x, t - h) = \varphi(x, t)$, $x > 0$. Then in (1) – (3) we get the problem:

$$u_t = a^2 u_{xx} + f(x, t, \varphi(x, t)), x > 0, h < t < 2h, \quad (4)$$

$$u(x, t)|_{t=h} = \varphi(x, h), x \geq 0, \quad (5)$$

$$u_x(0, t) - \gamma u(0, t) = \mu(t), t \geq h \quad (6)$$

with the conditions of agreed $\varphi_x(x, h) - \gamma\varphi(x, h) = \mu(h)$.

IV. SOLVING PROBLEM (4)-(6)

We are looking for a solution of problem (4) – (6) in the form of sum of three functions

$$u(x, t) = u_1(x, t) + u_2(x, t) + u_3(x, t), \quad (7)$$

where $u_i, 1 \leq i \leq 3$, respectively, take into account the influence only initial condition, the boundary condition and the inhomogeneity of the, that is, they are the solutions of such problems.

Problem 1. Find a function $u_1(x, t)$ that satisfies the conditions

$$\frac{\partial u(x, t)}{\partial t} = a^2 \frac{\partial^2 u(x, t)}{\partial x^2}, x > 0, h < t < 2h, \quad (8)$$

$$u(x, h) = \varphi(x, h), x \geq 0, \quad (9)$$

$$u_x(0, t) - \gamma u(0, t) = 0, h \leq t \leq 2h, \gamma > 0, \quad (10)$$

moreover, $\varphi_x(0, h) - \gamma\varphi(0, h) = 0$ is a condition of agreed.

Problem 2. Find a function $u_2(x, t)$ that satisfies equation (8) and conditions

$$u(x, h) = 0, x \geq 0, \quad (11)$$

$$u_x(0, t) - \gamma u(0, t) = \mu(t), h \leq t \leq 2h, \gamma > 0, \quad (12)$$

moreover, $\mu(h) = 0$ is a condition of agreed.

Problem 3. Find the function $u_3(x, t)$ that satisfies equation (1) and conditions (10), (11), which are agreed.

Solving problem 1

Let's expand the domain of definition of equation (8) and the initial condition to $x \in R$, $h \leq t \leq 2h$ and solve it with help of separation of variables method [6].

We obtain that

$$u_1(x, t) = \frac{1}{\sqrt{\pi(t-h)}} \int_0^\infty \varphi(x, h) \left\{ e^{-\frac{x^2 + \xi^2}{4a^2(t-h)}} ch \frac{x\xi}{2(t-h)} - \right. \\ \left. - \gamma \sqrt{\pi(t-h)} e^{\gamma^2(t-h) + \gamma(x+\xi)} \operatorname{erfc} \frac{x + \xi + 2\gamma(t-h)}{2\sqrt{t-h}} \right\} d\xi,$$

then $\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$, $x \geq 0$, $h < t < 2h$, $\operatorname{erfc} z = \frac{2}{\sqrt{\pi}} \int_0^z \exp\{-\xi^2\} d\xi$. Using the method of

mathematical induction, we prove that in case $x > 0$, $kh < t < (k+1)h$ the solution to problem 1 takes the form

$$u_1(x, t) = \frac{1}{\sqrt{\pi(t-kh)}} \int_0^\infty \varphi(x, kh) \left\{ e^{-\frac{x^2+\xi^2}{4a^2(t-kh)}} ch \frac{x\xi}{2(t-kh)} - \right. \\ \left. -\gamma\sqrt{\pi(t-kh)} e^{\gamma^2(t-kh)+\gamma(x+\xi)} \operatorname{erfc} \frac{x+\xi+2\gamma(t-kh)}{2\sqrt{t-kh}} \right\} d\xi. \quad (13)$$

Let's mark

$$G_1(x, y, t-kh) = \frac{1}{\sqrt{\pi(t-kh)}} e^{-\frac{x^2+y^2}{4a^2(t-kh)}} ch \frac{xy}{2a^2(t-kh)} - \\ -\gamma\sqrt{\pi(t-kh)} e^{-\gamma^2(t-kh)+\gamma(x+y)} \operatorname{erfc} \frac{x+y+2\gamma(t-kh)}{2\sqrt{t-kh}} \Bigg\}, \quad (14)$$

$x \geq 0, y > 0, kh < t < (k+1)h, k \in N$.

Definition. A function $G_1(x, y, t-kh)$ is called a Green's function of problem (1), (2), (3) if it satisfies the following conditions:

1) the function $G_1(x, y, t-kh)$ is continuous on x, y, t , continuously differentiable on t and twice continuously differentiable on x, y when $x > 0, y > 0, kh < t < (k+1)h, k \in N$, and possibly with the exception in the point $x = y, t = kh$;

2) the function $G_1(x, y, t-kh)$ by variables x and y satisfies the equation $\frac{\partial G_1}{\partial t} = a^2 \frac{\partial^2 G_1}{\partial x^2}$ everywhere except of the points $x = y, t = kh, k \in N$;

3) the function $G_1(x, y, t-kh)$ satisfies the boundary condition $(\partial / \partial x G_1(x, y, t-kh) - \gamma G_1(x, y, t-kh)) \Big|_{x=0} = 0$.

The Green's function satisfying this definition is constructed above and takes the form (14) and $G_1(x, y, t-kh) = G_1(y, x, t-kh)$.

Property of the solution of the problem 1

Theorem 1. If there exists a number $M > 0$ such that the initial function $\varphi(x, kh)$ is bounded when $x > 0, h > 0, k \in N, |\varphi(x, kh)| \leq M$, then the function $u_1(x, t)$ (13) when $x > 0, kh < t < (k+1)h$ is also bounded and the estimate

$$|u_1(t, x)| \leq M \left\{ \operatorname{erf} \frac{x}{2\sqrt{t-kh}} + e^{\gamma^2(t-kh)+\gamma x} \operatorname{erfc} \frac{x+2\gamma(t-kh)}{2\sqrt{t-kh}} \right\}$$

is true for it.

Solving the problems 2 and 3

Using the method [7], we get that

$$u_2(x, t) = \int_{kh}^t \mu(\tau) G_1(x, 0, t-\tau) d\tau, \quad (15)$$

$$u_3(x, t) = \int_{kh}^t d\tau \int_0^\infty f(y, \tau) G_1(x, y, t-kh) dy, \quad (16)$$

$x > 0, kh \leq t \leq (k+1)h, k \in N$.

Theorem 2. The solutions of problems 1, 2, 3 are determined by formulas (13), (15) and (16) respectively.

V. CONCLUSION

In this paper was formulated the third initial-boundary value problem for the heat conduction equation containing a nonlinear term dependent on the searched function with a deviation of the argument for the first time. For such equations, the initial condition is set on a certain interval. Physical and technical reasons for lateness can be transport delays, delays in information transmission, delays in decision-making, etc. The most natural are delays when modeling objects in ecology, medicine, population dynamics, etc. Features of the dynamics of vehicles in different environments (water, land, air) can also be taken into account by introducing a delay. Other physical and technical interpretations are also possible. The study of the molecular distribution of heat energy in various substances (solid bodies, liquids, etc.) leads to heat conduction equations. The Green's function of the third initial-boundary value problem is constructed for the nonlinear equation of heat conduction with a deviation of the argument, its properties are studied, and the formula for the solution is established.

REFERENCES

- [1] L.C. Evans, “*Partial Differential Equations*,” *Grad. Stud. Math.*, Amer. Math. Soc., Providence, RI, 2010, vol. 19, pp. 44-65. ISBN 978-0-8218-4974-3
- [2] Ian N. Sneddon, “*Elements of Partial Differential Equations*,” Dover Books of Mathematics, 2006, 352 p. ISBN - 10:9780486452975, ISBN – 13: 9780486452975.
- [3] A.B. Tayler, “*Mathematical Models in Applied Mechanics*,” Oxford, 2008, 288 p. ISBN – 10:0198515596
- [4] I. Kaur, Y. Mishin, W. Gust, “*Fundamentals of Grain and Interphase Boundary Diffusion*,” Chichester: John Wiley & Sons Ltd, 1995, 512 p.
- [5] J. Biazar and Z. Ayati, “An Approximation to the Solution of Parabolic Equation by Adomian Decomposition Method and Comparing the Result with Crank-Nicolson Method,” *International Mathematical Journal*, no. 39, 1925-1933, 2006.
- [6] Cole Kevin D., Beck James V., Haji-Sheikh A., Litkouhi Bahan, “*Heat conduction using Green's functions, Series in Computational and Physical Processes in Mechanics and Thermal Sciences (2nd ed.)*,” Boca Raton, FL: CRC Press, 2011. ISBN 978-1-43-981354-6
- [7] R.K.M. Thambaynagam, “*The Diffusion Handbook: Applied Solutions for Engineers*,” McGraw-Hill Professional, 2011. ISBN 978-0-07-175184-1
- [8] Y.M. Drin', I.I. Drin', S.S. Drin', Y.P. Stetsko, “The first boundary value problem for the nonlinear equation of heat conduction with deviation of the argument,” in *Proc. The 12th International Conference on Electronics, Communications and Computing*, 20-21 October, 2022, Chisinau, Republic of Moldova.
- [9] Y.M. Drin', I.I. Drin', R.Y. Drin', “The analytical view of solution of the second boundary value problem for the nonlinear equation of heat conduction with deviation of the argument,” in *Proc. The Eleven International Conference on “INFORMATICS AND COMPUTER TECHNIES PROBLEMS” (PICT – 2022)*, 10-13 October 2022, Chernivtsi, Ukraine, pp. 11-18.

Дослідження масштабування та розробка рекомендаційної системи методами глибинного навчання та технологій великих даних

Владислав Гой

*Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація—У даній статті представлено розробку механізму рекомендацій із використанням технологій великих даних і методів оптимізації, створених для прискорення обробки даних.

Ключові слова—*python, apache spark, sql, pandas, великі дані, обробка даних, рекомендаційна система, оптимізаційні методи.*

I. ВСТУП

Механізми або системи рекомендацій здебільшого використовуються для рекомендування відповідного та корисного вмісту потрібному користувачеві, щоб певним чином покращити загальний досвід користувача та надати йому нові функції. Такі механізми є потужними та широко використовуваними з точки зору використання величезної кількості даних і навчання розуміти переваги конкретних користувачів. Рекомендації допомагають користувачам легко переміщатися між мільйонами продуктів або тонами вмісту (статті/відео/фільми) і показують їм потрібну інформацію, яка може бути корисною для певної служби, якою вони користуються. Така функція допомагає знаходити інформацію від імені користувачів.

У більшості випадків користувачі вирішують, чи ефективний механізм щодо рекомендацій, і вони можуть вибрати вміст або ігнорувати його. Кожне рішення користувачів допомагає перенавчити модель машинного навчання за лаштунками на останніх даних, щоб мати можливість давати кращі рекомендації.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Багато дослідницьких робіт було проведено з використанням методів великих даних, включаючи інтелектуальний аналіз даних і методи машинного навчання. Застосувань цих технік дуже багато. Зокрема, механізм рекомендацій, який використовується в кількох сферах, таких як аналіз кошика (Amazon), соціальні мережі (LinkedIn, Twitter), уряд, освіта тощо.

Рекомендаційні моделі визначаються як системи, які створюють персоналізовані прогнози як результат. Ці системи застосовуються до електронної комерції, музики, фільмів, відео, книг, новин тощо. Залежно від способу створення рекомендацій їх поділяють на: системи рекомендацій на основі вмісту, системи рекомендацій спільної фільтрації та гібридні системи рекомендацій. Тим не менш, у всіх цих методах точність прогнозування цільового результату тісно пов'язана з кількістю інформації, отриманої від активних користувачів, для яких генеруються ці прогнози. У системі рекомендацій інтереси

користувачів, їхні переваги та антипатії використовуються для розробки профілю користувача, який використовується як фільтри. Створення точних профілів користувачів є важливим завданням, і точність продуктивності системи залежить від здатності навчених активних профілів споживачів представляти вподобання користувача. Початкова робота над рекомендаційними системами була з використанням спільної фільтрації, яка рекомендувала користувачам новинні статті та рекомендації музичних альбомів і виконавців із соціальної інформації. Після цього було багато робіт у сфері рекомендаційних систем, які допомагали споживачам ідентифікувати товари, послуги та контент, наприклад книги, дослідницькі статті, фільми, новини, роздрібний бізнес, цифрові продукти, споживчі товари тощо, застосовуючи різні алгоритми, які переглядає різних користувачів і предмети, щоб надати правильні пропозиції.

Лі та ін. [7] запропонував алгоритм для розпаралелювання частого виявлення набору елементів для пошуку прихованого шаблону для підтримки рекомендацій щодо запиту великого набору даних. Цей алгоритм дає змогу зменшити вартість обчислень шляхом розподілу обчислень між вузлами кластера таким чином, що кожен вузол виконує незалежний набір завдань видобутку. Вони досягли найкращої продуктивності завдяки розподілу обробки за допомогою інфраструктури MapReduce на кластер комп'ютерів.

Наджафабаді та інші., [2] представили глибоке навчання — це найшвидший сегмент машинного навчання, який використовує багато рівнів глибоких нейронних мереж для вивчення рівнів представлення та абстракції, що робить сенс даних, зазвичай за допомогою штучних нейронних мереж. Глибоке навчання сприяє значному прогресу в галузях, підприємствах та електронній комерції.

Багато систем рекомендацій для соціальних мереж було розроблено з використанням взаємодії та поведінки користувачів. У [3] запропоновано підхід до аналізу інтересу користувачів до платформи twitter. Вони поєднали аналіз настроїв і класифікацію твітів, аналізуючи теми, які обговорювали користувачі. Реалізація їх роботи дає обнадійливі результати.

III. МЕТОДИ

У рамках цієї статті систему рекомендацій було реалізовано за допомогою Python Apache Spark ML Pipeline, реалізація альтернативних найменших квадратів. Модель приймає навчальний набір даних як кадр даних (`pyspark.sql.DataFrame`) і кілька параметрів, які керують процесом створення моделі [3].

Користувачі можуть легко сформулювати мету своїх запитів за допомогою Spark SQL, тоді як Spark виконує більш складний процес оптимізації запитів [4]. З моменту введення визначених користувачем функцій `pandas` у `spark 2.3` [3] користувачі можуть визначати власні `python`-функції, які можна запускати пакетами, що дає їм гнучкість, їм потрібно створювати запити, які стосуються надзвичайно спеціалізованих випадків використання.

Кожен ідентифікатор моделі, пара ідентифікаторів і майбутній новий рядок повинні бути перевірені Spark SQL. Замість виконання грубого перехресного з'єднання Spark ми можемо просто перебрати всі рядки та пам'ять за допомогою UDF. Виявляється, для певного набору даних така програмна реалізація може майже подвоїти швидкість обробки.

Ідея інвертованих індексів також дуже корисна для оптимізації проблеми. Основний принцип цієї концепції полягає в тому, що кожен рядок сховища функцій містить лише невелику кількість відмінних функцій. Кожен рядок набору даних містить від 10 до 20 характеристик. Тому було б дуже корисно, якби ми змогли знайти техніку для повторення рядків і повернення відповідей, використовуючи 0 функцій на час рядка на відміну від 0 часу

фільтрів. Інвертовані індекси є методом для досягнення цього. Однак загальне уявлення полягає в тому, що нам потрібен алгоритм, який масштабується пропорційно до того, наскільки розрідженими є дані.

Існують альтернативні бібліотеки Python, такі як `itertools`, які призначені для того, щоб програми Python працювали набагато швидше [1]. `Pandas` зазвичай налаштовано на простоту використання, а не на продуктивність.

Як виявилось, потрібно використовувати `itertools` для нашого випадку використання. `Reduce` в `NumPy` може бути гідним варіантом для розгляду для інших UDF [3], але це не стосується нашого випадку використання.

Тому можна зробити це в 2,5 рази швидше, ніж наш попередній підхід, використовуючи цей модуль Python, який оптимізований для виконання дуже конкретних речей [2]. Програмна реалізація оптимізаційного методу з використанням `itertools` дає результат, який близько у 100 разів швидший, ніж нативне рішення `Spark SQL`.

```
def udf_wrapper(feature_data: pd.DataFrame) -> pd.DataFrame:
    for ids in itertools.groupby(feature_data['feature_ids'], key=lambda x: x[0]):
        valid_ids = set()
        for _, id_ in ids:
            included = set(
                itertools.chain.from_iterable(
                    [incl_ids[f_id] for f_id in id_ if id_ in incl_ids]
                )
            )
            excluded = set(
                itertools.chain.from_iterable(
                    [excl_ids[f_id] for f_id in id_ if id_ in incl_ids]
                )
            )
            filtered = incl_ids.difference(excl_ids)
            valid_ids = valid_ids.union(filtered)

        for f_id in valid_ids: counter[f_id] += 1

    return pd.DataFrame([[id_, count] for id_, count in counter.items()])

groupby_udf = pandas_udf(
    f=udf_wrapper,
    returnType=StructType(
        [
            StructField('id', IntegerType()),
            StructField('batchCount', IntegerType())
        ]
    )
)
```

Рис. 1 Програмна реалізація оптимізаційного методу за допомогою `itertools`

IV. РЕЗУЛЬТАТИ

Результатом дослідження є реалізовані оптимізаційні методи для обробки даних для середовища Apache Spark, а саме: метод інвертованих індексів та за допомогою використання нативного Python (itertools).

Для всіх експериментів використовувалися мова програмування Python 3.10, PyCharm і MacBook Pro з операційною системою Ventura 13.0, процесором Apple M1 Pro 16 ГБ оперативної пам'яті

Оцінивши запропоновані методи, можна сформулювати порівняльну характеристику швидкодії існуючих рішень та розроблених, зображену на рисунку 2.

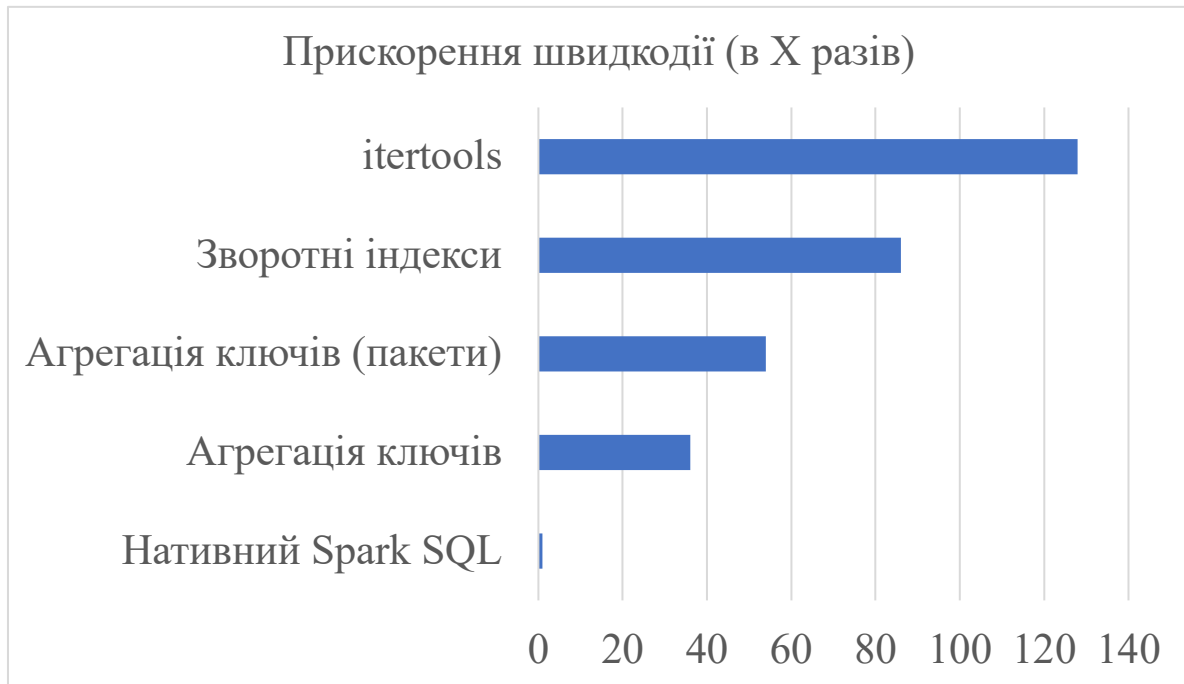


Рис. 2 Порівняння швидкодії методів

Як зображено на рисунку, дані методи дають змогу обробляти дані в рази швидше, порівняно із звичайним підходом в Spark SQL, а саме метод зворотних індексів прискорює обробку даних у 86 разів, а метод за допомогою використання itertools в 128 разів.

V. ОБГОВОРЕННЯ

Запропоновані підходи для оптимізації обробки даних в межах рекомендаційної системи здатні значною мірою прискорити існуючі та нові методи обробки даних не тільки в рекомендаційних системах, а й в інших системах, які використовують обробку великих даних та моделі машинного навчання.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Запропоновані підходи є перспективними кандидатами на подальші вдосконалення, а саме їх дослідження в інших сферах машинного навчання та великих даних, оскільки можуть принести вагомий внесок та прискорити різного роду системи та алгоритми, які ними використовуються.

Отримані багатообіцяючі результати є лише відправною точкою. Можна розглянути можливість розгортання цієї реалізації на багатовузловому кластері, щоб оцінити його

масштабованість, продуктивність (зокрема час обчислень) і точність у розподіленому режимі.

Інші дослідницькі зусилля слід присвятити експерименту з іншими інструментами та фреймворком великих даних, таким як Apache Mahout, і порівняти його продуктивність з Apache Spark.

VII. ВИСНОВКИ

Системи рекомендацій були давно розроблені та інтегровані на багатьох веб-сайтах, особливо на веб-сайтах електронної комерції, давно. Вони довели свою потужність у наданні персоналізованого, налаштованого веб-вмісту різним користувачам, рекомендуючи вміст (тобто елементи у випадку веб-сайту електронної комерції), що цікавить кожного користувача, і таким чином пом'якшує проблему перевантаження інформацією для користувача.

Для систем рекомендацій існують різні техніки та підходи. Одним із широко використовуваних методів є спільне фільтрування, яке збирає записи про взаємодію між користувачами та товарами, історію покупок, щоб визначити смак користувача та таким чином рекомендувати товари, які відповідають його смаку.

Ця дослідницька робота робить внесок у сучасні рекомендаційні системи в тому сенсі, що забезпечує реалізацію великомасштабної рекомендаційної системи на основі Apache Spark. Це стало результатом інтенсивного вивчення літератури, а також проведення багатьох експериментів з використанням мови програмування Python на основі Apache Hadoop і Spark. Дослідження охоплювало кілька тем, а саме: феномен великих даних, різні техніки та підходи до рекомендаційних систем разом із їхніми плюсами та мінусами, виклики, пов'язані з великими даними щодо рекомендаційних.

Ця робота вирішила проблему масштабованості та оптимізації обробки даних шляхом використання реалізованих оптимізаційних методів на основі розробленої рекомендаційної системи.

VIII. КОНФЛІКТИ ІНТЕРЕСІВ

Автор зазначає, що дослідження проводилося за відсутності будь-яких комерційних чи фінансових відносин, які можна було б витлумачити як потенційний конфлікт інтересів.

ЛІТЕРАТУРА

- [1] Percival H.J.W., Gregory B., Architecture Patterns with Python Enabling Test- Driven Development, Domain-Driven Design, and Event-Driven Microservices, O'Reilly Media, 2020, p .304.
- [2] Najafabadi, Maryam & Villanustre, Flavio & Khoshgoftaar, Taghi & Seliya, Naeem & Wald, Randall & Muharemagic, Edin, Deep learning applications and challenges in big data analytics, Journal of Big Data, 2015, pp. 34-45.
- [3] Mengle S., Gurmendez M., Mastering Machine Learning on AWS Advanced Machine Learning in Python Using SageMaker, Apache Spark, and TensorFlow, Packt Publishing, 2019, p. 306.
- [4] Ramalho L., Fluent Python Clear, Concise, and Effective Programming, O'Reilly Media, 2015, p. 790.
- [5] Chambers B., Zaharia M., Spark The Definitive Guide: Big Data Processing Made Simple, O'Reilly Media, 2018, p. 603.
- [6] Li H, Wang Y, Zhang D, Zhang M, Chang EY, Pfp parallel fp-growth for query recommendation, In Proceedings of the 2008 ACM conference on recommender systems, RecSys'08, Lausanne, ACM Press, 2008, p. 107.
- [7] Sandy Ryza, Advanced Analytics with PySpark, Patterns for Learning from Data at Scale Using Python and Spark, 2022, pp. 120-168.
- [8] Romeo Kienzler, Md. Rezaul Karim, Sridhar Alla, Siamak Amirghodsi, Apache Spark 2: Data Processing and Real-Time Analytics, 2018, pp 20-88.

Дослідження шкідливого програмного забезпечення

Юрчук Дмитро, Савка Іван

*Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна*

Анотація—В даному дослідженні розглядаються можливі методи для аналізу та виявлення шкідливого програмного забезпечення, подаються необхідні для цього засоби, описується методика для динамічного аналізу.

Ключові слова—*Безпека Даних, Шкідливе Програмне Забезпечення, Віруси, Статичний Аналіз, Динамічний Аналіз, Модель.*

I. ВСТУП

Шкідливе програмне забезпечення (ПЗ) негативно впливає не тільки на інформаційну індустрію, а й на різні сфери людської діяльності, оскільки у сучасному суспільстві майже у кожній галузі людина безпосередньо взаємодіє з електронними пристроями. Вірусні програми можуть передаватися різними шляхами, наприклад, через фізичні носії даних, електронну пошту, мережу чи маскування під звичайні користувацькі програми [1]. Оскільки будь-яка сфера маніпулює з даними, то відповідно виникає загроза втрати цих даних або несанкціонованого доступу до них. Для цього необхідно використовувати дієві інструменти, щоб превентивно захиститися від комп'ютерних загроз. На даний час існує велика кількість засобів, які виявляють віруси – від готових антивірусних програм до динамічного аналізу та візуалізації вірусів [2]. Проте виникають нові види шкідливого ПЗ, інші щоразу оновлюються, відповідно не всі програмні продукти можуть ефективно виявляти такі загрози.

II. АНАЛІЗ ОСТАННІХ ПУБЛІКАЦІЙ

Серед основних методів виявлення шкідливого ПЗ можна виділити сигнатурний, евристичний та аномальний (поведінковий) [3, 4]. Сигнатура – це алгоритм або хеш, який унікально ідентифікує конкретне зловмисне ПЗ, а евристика – це набір правил, визначених після аналізу поведінки зловмисного ПЗ. Часто буває, що сигнатурні методи не можуть виявити невідомі зловмисні програми, вони потребують великої кількості робочої сили, часу та грошей для вилучення унікальних підписів. Крім того, ще одним недоліком є нездатність протистояти зловмисним програмам, які змінюють свої коди. На відміну від сигнатурного, евристичний метод, який зазвичай ґрунтується на ймовірнісних алгоритмах, дозволяє виявляти раніше невідомі віруси. Ці підходи вимагають аналізу зловмисного ПЗ перед визначенням правил і евристик.

Відповідно аналіз ПЗ можна здійснювати двома основними методами [5]:

- статичним – коли проводиться аналіз файлу з вірусом, не виконуючи його (передбачає дослідження заголовків, вбудованих ресурсів, метаданих, хешу);
- динамічним – відбувається запуск шкідливого ПЗ та слідкування за поведінкою того чи іншого вірусу (важливо не запускати шкідливе ПЗ на реальному пристрої, оскільки це реальні віруси, які можуть створити загрозу вашому програмному забезпеченню).

Використовуючи динамічний аналіз шкідливого ПЗ, програма виконується в ізолюваному середовищі та відстежується її поведінка (наприклад, системні виклики API). Динамічні методи аналізу є більш ефективними порівняно з статичними, вони застосовуються

проти обфускації коду та нещодавно випущених шкідливих програм («нульового дня»). Крім того, виділяють гібридний метод, який поєднує в собі статичний та динамічний методи, що дозволяє ще ефективніше виявити шкідливий код.

III. МОДЕЛЬ РОЗПІЗНАВАННЯ ШКІДЛИВОГО ПЗ

Загалом, динамічна модель відображає часові характеристики системи та послідовність взаємодії функцій у часі, оперуючи такими поняттями, як стан системи, події, перехід із одного стану в інший, умови переходу, послідовність подій. Динамічний підхід до аналізу шкідливого ПЗ є досить ефективним, оскільки дозволяє почергово запускати зразки в деякому ізолюваному середовищі, детально досліджувати поведінку вірусу та збирати різноманітні дані у вигляді логів [6].

Алгоритм динамічного дослідження шкідливого ПЗ наведено на Рисунку 1.



Рисунок 1 – Алгоритм динамічного аналізу шкідливого ПЗ.

При цьому добре побудована модель на основі цих даних може з великою ймовірністю визначати шкідливе ПЗ.

Для аналізу поведінки вірусів можна скористатися пісочницею Cuckoo Sandbox, яка працює в ізолюваному середовищі. Даний інструмент дозволяє повністю автоматизувати аналіз, однак для цього необхідно встановити «гостьовий агент» у середовище. Інколи зловмисні програми можуть маскуватися або змінювати свою поведінку в цьому середовищі. Подібно до Sandbox технології існує Virtual Machine Introspection (VMI), яка дозволяє користувачеві контролювати виконання процесу, файлові операції, відстежувати системні виклики, функції ядра (дозволяє простежити поведінку руткітів). Оскільки на цій віртуальній машині не встановлюється жодних агентів, шкідливі програми демонструють свою

природню поведінку, що дозволяє краще зробити їх аналіз. Для дослідження методом VMI можна скористатися інструментом DRAKVUF [7].

В якості бази вірусів можна скористатись відкритими ресурсами: GitHub репозиторіями «theZoo» [8] і «api_sequences_malware_datasets» [9], датасетом CIC-MalMem-2022 [10] тощо. Оскільки вірусне ПЗ стрімко розширюється, то до існуючих даних варто додати власні проаналізовані зразки, щоб покращити датасет для побудови моделей. Далі на основі отриманих даних та їх обробки натренувати вибрані моделі, які зможуть розпізнавати шкідливе ПЗ, та оцінити їх точність.

IV. РЕЗУЛЬТАТИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В даній роботі було досліджено методи аналізу та розпізнавання шкідливого ПЗ, розглянуто інструменти для досліджень, їх переваги та недоліки. Крім того, описано методику, за допомогою якої можна здійснювати динамічний аналіз шкідливого ПЗ та його виявлення. В подальшому розробка ефективної моделі дасть можливість визначати не тільки вже існуючі види шкідливого програмного забезпечення, а й досить добре розпізнавати нові види вірусів.

V. ВИСНОВКИ

Безпека інформації є важливою, оскільки будь-які вірусні атаки несуть за собою в першу чергу економічні збитки для будь-якої організації. Аналіз вмісту програм, оновлення наборів даних шкідливого ПЗ і використання методів штучного інтелекту дозволяє виявляти нові загрози, розробляти та впроваджувати ефективні заходи для забезпечення потрібного рівня кіберзахисту.

ЛІТЕРАТУРА

- [1] Reiner Creutzburg, Wikipedians, *Handbook of Malware 2016 - A Wikipedia Book*. Germany: Wikipedia, 2016, doi: 10.13140/RG.2.1.5039.5122
- [2] M. J. Hossain Faruk et al., “Malware Detection and Prevention using Artificial Intelligence Techniques,” *2021 IEEE International Conference on Big Data (Big Data)*, 2021, pp. 5369-5377, doi: 10.1109/BigData52589.2021.9671434.
- [3] Z. Bazrafshan, H. Hashemi, S. M. Hazrati Fard, and A. Hamzeh, “A survey on heuristic malware detection techniques,” *IKT 2013 - 2013 5th Conference on Information and Knowledge Technology*, 2013. pp. 113-120, doi: 10.1109/IKT.2013.6620049.
- [4] Kurt Baker. “Malware analysis.” CrowdStrike.com. <https://www.crowdstrike.com/cybersecurity-101/malware/malware-analysis> (accessed Nov. 15, 2022).
- [5] Daniel Gibert, Carles Mateu, Jordi Planes, “The rise of machine learning for detection and classification of malware: Research developments, trends and challenges,” *Journal of Network and Computer Applications*, vol. 153, issue C, 2020, doi: 10.1016/j.jnca.2019.102526.
- [6] .A. Melvin, and G.J. Kathrine, “A Quest for Best: A Detailed Comparison Between Drakvuf-VMI-Based and Cuckoo Sandbox-Based Technique for Dynamic Malware Analysis,” in *Intelligence in Big Data Technologies—Beyond the Hype*, J. D. Peter et al. (eds.), Singapore, 2021, pp. 275-290, doi: 10.1007/978-981-15-5285-4_27.
- [7] A Live Malware Repository, GitHub.com, Nov. 15, 2022. [Online]. Available: <https://github.com/ytisf/theZoo>.
- [8] API sequences malware datasets, GitHub.com, Nov. 15, 2022. [Online]. Available: https://github.com/khas-ccip/api_sequences_malware_datasets.
- [9] Malware Memory Analysis, Canadian Institute for Cybersecurity, Nov. 15, 2022. [Online]. Available: <https://www.unb.ca/cic/datasets/malmem-2022.html>.

Напрямки удосконалення міжнародних стандартів у сфері оцінки якості програмних продуктів ISO/IEC 9126 та ISO/IEC 14598

Микола Кузь¹, Ігор Лазарович¹, Руслан Мішагін²

¹Кафедра інформаційних технологій
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна

²Лицей імені Романа Шухевича
Івано-Франківськ, Україна

Анотація—Здійснено аналіз вимог та сфери застосування міжнародних стандартів у сфері оцінки якості програмних продуктів серій ISO/IEC 9126 та ISO/IEC 14598. Виявлено проблемні питання, які унеможливають повноцінне застосування цих стандартів для оцінки якості програмних продуктів із-за відсутності в них необхідних методів та методик. Розроблено проект стандарту, який урегулює ці питання, а методики, наведені в цьому проекті є повноцінним доповненням вище наведених серій стандартів.

Ключові слова—Стандарт, Якість програмних продуктів, Методика.

I. ВСТУП

Питання якості продукції, зокрема і програмного забезпечення, є актуальним завжди, як з точки зору споживачів, так і виробників продукції. Вимоги до якості програмних продуктів регламентуються міжнародними стандартами серії ISO/IEC 9126 [1-4] та ISO/IEC 14598 [5-10], які також прийняті і в Україні як національні стандарти. Сфера охоплення цими нормативними документами процесів оцінювання якості програмних продуктів, за даними з [1], зображена на рис. 1.

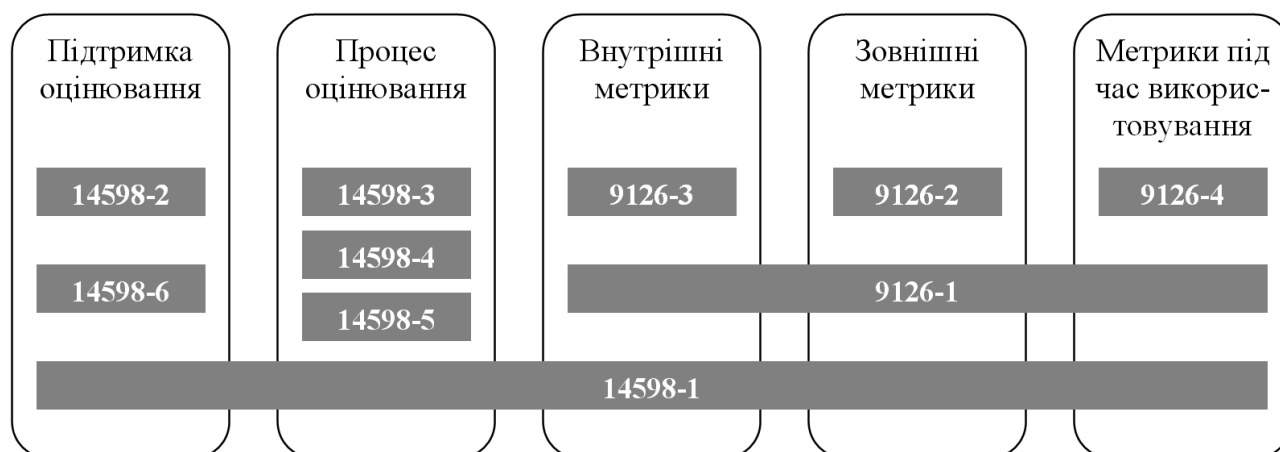


Рисунок 1 – Сфера дії стандартів серії ISO/IEC 9126 та ISO/IEC 14598

Нормативні документи [1-10] чітко розмежовують функції та повноваження розробників, зовнішніх оцінювачів та користувачів програмних продуктів, розмежовують сфери застосування внутрішніх, зовнішніх метрик якості програмних продуктів та метрик якості під час використання.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Більш детальний аналіз стандартів [1-10] засвідчив наявність проблемних питань, що стосуються застосування цих документів для оцінки якості програмних продуктів. Зокрема, стандарти серії ISO/IEC 9126 [1-4] містять номенклатуру атрибутів якості програмних продуктів, наводять методи визначення тих чи інших атрибутів. Атрибути, в свою чергу, погруповані в метрики підхарактеристик якості, які є складовими характеристик якості. Ці стандарти не містять жодних методик для визначення показників якості, за винятком методів визначення атрибутів якості. Методи визначення якісних показників підхарактеристик, характеристик якості та узагальненого показника якості в цих документах відсутні. Відсутні і методи формування вагових коефіцієнтів для підхарактеристик і характеристик якості програмних продуктів.

Не вирішують цієї проблеми і стандарти серії ISO/IEC 14598 [5-10]. Вони тільки регламентують процедури оцінки якості, описують певні підходи з організації робіт щодо оцінки якості на різних стадіях життєвого циклу програмних продуктів. І якщо і наводять деякі рекомендаційні речі, то використати їх в практиці неможливо. Для прикладу, в них наводиться шкала оцінки програмних продуктів, яка не має оцифрованих поділок, а лише розбивається на діапазони допустимих і недопустимих значень.

Таким чином, кожен розробник програмного забезпечення сам повинен розробляти певні методики для реалізації процедур оцінки програмних продуктів з метою застосування стандартів [1-10]. А це, в свою чергу, буде порушенням принципу уніфікації продукції. Результати оцінки якості програмних продуктів, отримані з використанням різних методик, не дадуть змоги порівняння якісних параметрів програмних продуктів однотипного призначення.

III. МЕТОДИ

Найбільш оптимальним способом вирішення проблеми уніфікації методів оцінки якості програмних продуктів є розроблення єдиного нормативного документу, який би регламентував застосування вимог всіх чинних міжнародних стандартів у сфері якості програмних продуктів [1-10]. Такий вид стандартів досить часто застосовується за потреби об'єднати і доповнити чи конкретизувати вимоги певної серії нормативних документів.

Структура такого виду стандартів формується наступним чином:

- у розроблюваному стандарті наводиться цитата певної вимоги (яка не в повній мірі розкрита), наприклад, першого із серії стандартів;
- слідом за цитатою наводиться розроблена методика, що в повній мірі реалізує дану вимогу чи деталізує її.

Така процедура застосовується до всіх наступних стандартів із серії. При цьому, послідовність викладення матеріалів у розроблюваному стандарті повинна чітко відповідати вимогам до побудови нормативних документів.

Розроблюваний стандарт не замінює серію стандартів, а тільки уточнює і конкретизує їхні вимоги, містить методики реалізації цих вимог.

Такий стандарт повинен пройти повну процедуру погодження та реєстрації як і будь-який інший нормативний документ такого типу. Тільки в такому випадку дія цього стандарту буде рівнозначною із стандартами, вимоги яких цей документ конкретизує. Крім того рівень стандарту також має бути рівнозначний (міжнародним, у випадку удосконалення міжнародних стандартів чи національним – для національних стандартів).

IV. РЕЗУЛЬТАТИ

З метою ефективного застосування стандартів [1-10] було розроблено проект стандарту, структурна модель якого наведена на рис. 2.

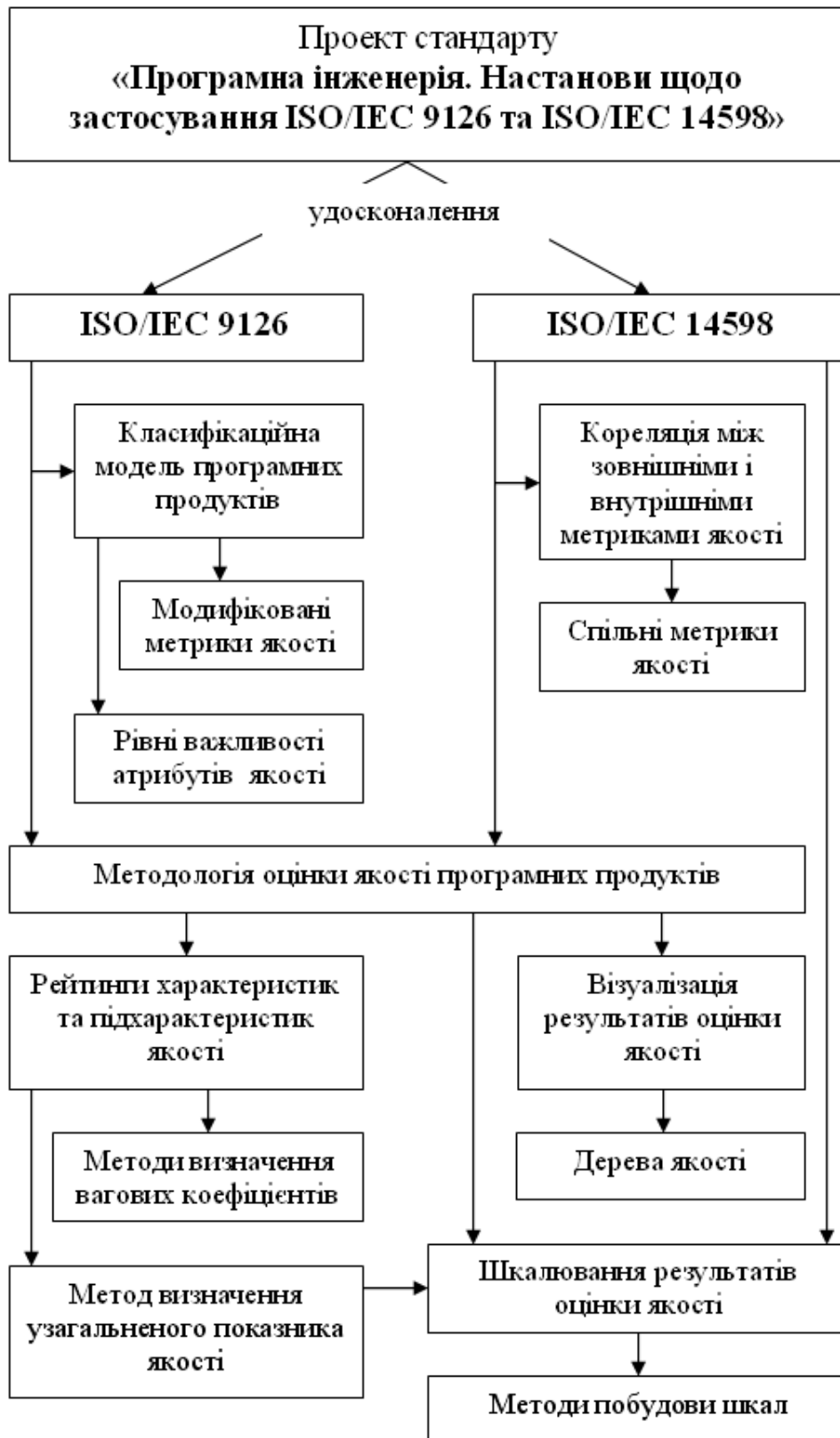


Рисунок 2 – Структурна модель проекту стандарту, що регламентує застосування стандартів серії ISO/IEC 9126 та ISO/IEC 14598

Спільним удосконаленням обох серій стандартів є розроблена методологія оцінки якості програмних продуктів. Вона містить методіку рейтингування характеристик та підхарактеристик якості програмних продуктів. Складовими цієї методіки є розроблені два методи визначення вагових коефіцієнтів, а також метод визначення узагальненого показника якості програмних продуктів. Останній дає змогу порівняти якісні показники двох чи кількох однотипних програмних продуктів. Методологія також містить методи візуалізації результатів оцінки якості програмних продуктів, що графічно відображаються у вигляді дерев якості та методи шкалювання цих результатів, зокрема підібрана найбільш оптимальна шкала, побудована на основі ряду Фібоначчі.

Удосконаленням стандартів серії ISO/IEC 9126 [1-4] є розроблена класифікаційна модель програмних продуктів, яка дозволяє здійснювати модифікацію метрик якості з врахуванням рівнів важливості атрибутів якості в цих метриках.

Що стосується удосконалення стандартів серії ISO/IEC 14598 [5-10], то вдалось встановити кореляцію між зовнішніми та внутрішніми метриками якості, завдяки чому сформована номенклатура спільних метрик якості. Такий підхід дозволяє на стадії оцінки внутрішньої якості програмних продуктів визначити показники якості спільних метрик і, вже на стадії оцінки зовнішньої якості, вдруге їх не визначати.

V. ОБГОВОРЕННЯ

Подальше затвердження розробленого стандарту і надання йому чинності дасть змогу в повній мірі застосовувати стандарти [1-10] для оцінки якості програмних продуктів, визначення узагальнених показників якості, порівняння значень яких між собою дозволить визначити переваги в якості того чи іншого програмного продукту.

VI. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Наступним етапом є експериментальна апробація розробленого стандарту та, за необхідності, удосконалення методів оцінки, наведених в ньому.

VII. ВИСНОВКИ

Розроблено проекту стандарт, що регламентує застосування стандартів серії ISO/IEC 9126 та ISO/IEC 14598. Використання розробленого стандарту дасть можливість реалізувати всі вимоги стандартів [1-10] для оцінки якості програмних продуктів на всіх етапах життєвого циклу. Стандарт буде корисним інструментом для розробників програмного забезпечення, незалежних оцінювачів та користувачів програмних продуктів.

ЛІТЕРАТУРА

- [1] *Software engineering — Product quality — Part 1: Quality model*, ISO/IEC 9126-1,2001
- [2] *Software engineering — Product quality — Part 2: External metrics*, ISO/IEC TR 9126-2,2003
- [3] *Software engineering — Product quality — Part 3: Internal metrics*, ISO/IEC TR 9126-3,2003
- [4] *Software engineering — Product quality — Part 4: Quality in use metrics*, ISO/IEC TR 9126-4,2004
- [5] *Information technology — Software product evaluation — Part 1: General overview*, ISO/IEC 14598-1,1999
- [6] *Software engineering — Product evaluation — Part 2: Planning and management*, ISO/IEC 14598-2,2000
- [7] *Information technology — Software product evaluation — Part 3: Process for developers*, ISO/IEC 14598-3,2000
- [8] *Software engineering — Product evaluation — Part 4: Process for acquirers*, ISO/IEC 14598-4,1999
- [9] *Information technology — Software product evaluation — Part 5: Process for evaluators*, ISO/IEC 14598-5,1998
- [10] *Software engineering — Product evaluation — Part 6: Documentation of evaluation modules*, ISO/IEC 14598-6,2001

Analysis of methods and existing software implementations for automatic resume generation

Vasyl Ivanochko

*Department of Information Technology
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine*

Abstract—This research is intended for evaluation and analysis of existing programs that allow to automatically generate a resume based on the data entered by the user for subsequent submission as a response to a vacancy. Criteria will be formed that will check the program for compliance with standards of use and convenience for evaluation in the research.

Keywords—Research, Resume, CV, Analysis, Candidate, Vacancy, Recruitment.

I. INTRODUCTION

A resume is very often called CV. CV is an abbreviation of the Latin word "curriculum vitae" which translates as "life path" [4]. A professional resume gives a brief overview and a good idea of a person's life. The major information about experience, education, skills, completed courses, etc., is placed in the resume. That is all the information that can be useful for a potential employer.

That is why CVs are extremely popular. Now, in fact, for every vacancy, recruiters require a resume file sent in response from the candidate. However, not everyone wants to do it, spends a lot of time on it, and not everyone knows how to do it correctly. The ability to export DOC or PDF requires a paid membership. The monthly subscription fee for this program is \$25.

Given the aforementioned information, several programs allow you to create your CV using the given data. We shall now discuss them.

I carried out this study to fairly assess and contrast the current applications of this subject. The outcomes of which will be the standards for assessing programs on this subject and putting them to use in actual implementations.

II. ANALYSIS OF LATEST RESEARCH AND PUBLICATIONS

The past research is about methods of creating or generating CVs from some external tools. For example, in the "Automatic Generation of CVs from Online Social Networks"[1] research about a method of creating a tool for the generation of resumes from the information that is placed on public storage, such as social networks.

Another research mentioned that existing templates for automatic generation of resume is very expensive. "It has been calculated that the generation of a CV takes an average 3-hour period and that an automated system could save at least \$25,000 per 100 generated CV's." wrote in the "Intelligent Agents and Apache Cocoon for a CV Generation System,"[2]. Also here is mentioned that existing implementation are hard to use for non-developer user. "For example, it is hard to imagine a standard Philosophy Ph.D. student typing an XML file and applying a XSL transformation to it" [2].

Both publications talking that this technology is a very in demand because of rapidly developing technologies. So, I decided to create a criteria for evaluating an implementation for automation resume generation.

III. CRITERIA AND THEIR DESCRIPTION

- *Ability to add and edit previous work experience*

Employers undoubtedly pay the greatest attention to this characteristic than any other. Of course, this choice will be available on every application, but it is still important to consider the focus on the resume, the possibility of including a description, and the order of the list of prior experience components. It is well recognized that the order of the list should be established from most pertinent to oldest positions.

- *The presence of tips for the correct resume for the user*

Not every applicant is adept at properly structuring a resume and adding information to it.

- *Ability to add and customize skills*

The second-most crucial component of the resume, from which the interviewer might infer information about the applicant during the CV review stage. It's critical that the software accurately depicts a potential employee's degree of skill proficiency in addition to simply confirming that talent is there on their résumé.

- *Adding education data*

Education is a crucial factor for the organization since, frequently, a person might enroll in standard courses and regard himself as an expert without developing skills like the capacity for research, cooperation, gaining additional information in numerous sectors, and broadening his views. The order used here must be analogous to experience, going from more recent to older. Additionally, consider if the reader is paying attention to the expertise rather than the location or date of gaining the specialism.

- *Adding data about courses and certifications*

Outside of academic institutions, a wide variety of courses, educational programs, and certificates are offered.

- *Add and customize a short description about yourself*

This area allows the applicant to quickly introduce himself, explain who he is, why he is.

- *The section for adding information about which languages the candidate speaks*

Knowledge of languages like German, French, or another is a huge bonus when deciding between two applicants because it frequently occurs that the client on the project is a foreigner (not necessarily an English-speaking one).

- *Adding a photo and the ability to hide it*

It is commonly known that to prevent acts of racism, sexism, or other discrimination against the applicant, it is illegal in several nations to include a photo, gender, or marital status on a resume.

- *Adding link data and contact methods*

Various communication techniques could be employed in various nations, or even in the same. It's also crucial to realize that various occupations could call for references to various online accounts. Here, it's important to consider both the types of communication that may be added to the CV in the software and how they are presented on the resume.

- *Resume templates*

The option to pick from the program's current templates the one you like most and the one is best suited for the chosen position. The diversity and quantity of templates offered by the application.

- *Choosing the main color of the resume*

Not essential, but a pleasant aspect of the requirements. The option to select the color that will serve as the resume's primary hue. This argument is significant for businesses who produce staff resumes for clients using a conveyor-style template. The hue must also match the company's color scheme, which is crucial.

- *Adding personalized sections to your resume*

It is impossible to predict which portions and pieces of information the hiring team will find most crucial regarding an applicant.

- *Section editing*

Programs may refer to the same sections differently. Sometimes it's vital to refer to a section as "Employment History" rather than "Experience," for instance. Additionally, it is essential to be able to manually modify the section name in the relevant language if the application lacks the option to alter the language.

- *The ability to preview resumes*

The user must view the resume as it is created to comprehend how it will seem with the entered data. Additionally, it will prevent any minor or grammatical problems when the template is filled out with data.

- *Confirmation of data deletion actions*

Dialog boxes should be provided to confirm the user's deletion of information to prevent circumstances where the user mistakenly deletes a portion or some of data.

- *Saving entered data and created templates*

The program must be able to save the created resumes, edit them as needed, and export them for the user's data to remain intact.

- *Program usage plan*

Most of the programs on this subject are paid, therefore you must routinely make payments to utilize them. This criterion assesses whether features may be accessed for free and how much it will cost to utilize the application to get the core set of capabilities.

IV. ANALYSIS OF EXISTING SOLUTIONS

The following programs participated in my research:

1. Resume.io [3]
2. cvmaker [4]
3. LinkedIn [5]

All of them are quite similar, and have extremely similar functionality, but they differ in their features. And they are quite different according to the criteria I described.

Let's start with the first program. This is **resume.io**. The program is very easy, convenient, and easy to use. Works quickly and without interruptions. On the user side, everything is easy, uncluttered, and neat.

Table 1. Analysis of **resume.io** by criteria

Criteria	Description
Ability to add and edit previous work experience	The program provides such an opportunity. You can briefly explain the position, the location of work, the time spent in this role, and the requirements and tasks of this employment. When adding a new item, it is not inserted in the typical sequence, however, this may be adjusted manually.
The presence of tips for the correct resume for the user	Tips and recommendations guide the user through the whole resume-creation procedure.
Ability to add and customize skills	It is possible to add skills from the program's existing collection as well as your own. It is also possible to specify the degree of mastery for each talent and, if required, suppress the display of this choice for a certain ability.
Adding education data	The criteria listed above operate flawlessly and fulfill resume management requirements.
Adding data about courses and certifications	The criteria listed above operate flawlessly and fulfill resume management requirements.
Add and customize a short description about yourself	The criteria listed above operate flawlessly and fulfill resume management requirements.
The section for adding information about which languages the candidate speaks	The criteria listed above operate flawlessly and fulfill resume management requirements.
Adding a photo and the ability to hide it	If a photo is included in the template, you can add or delete it, and the template will automatically alter it to match the style.
Adding link data and contact methods	You can include multiple links with personalized signatures, but they are presented as clickable links in the CV, making it impossible to see the link, for example, in print.
Resume templates	About 30 templates are available, with some allowing you to export resumes in PDF and DOC formats both.
Choosing the main color of the resume	The resume color can be set from the complete RGB palette.
Adding personalized sections to your resume	There are various configurable parts, such as "Hobbies" or "Internships," as well as an entirely "Customizable Section."
Section editing	Each area is independently configurable, and section titles can be changed.
The ability to preview resumes	View the resume while working and do a rapid update after adding fresh data.
Confirmation of data deletion actions	A confirmation dialog is displayed before any deletion, a confirmation dialog.
Saving entered data and created templates	Data is promptly stored upon the entry of new values.
Program usage plan	The ability to export DOC or PDF requires a paid membership. The monthly subscription fee for this program is \$25.

The next program from the list is **cvmaker**. The application is designed in the form of a tour. You fill in your data step by step. You can easily return to the previous step without losing data from the current one.

Table 2. Analysis of **cvmaker** by criteria

Criteria	Description
Ability to add and edit previous work experience	You can briefly define the duties and responsibilities of this post as well as the role's responsibilities, working environment, and time commitment.
Adding education data	
Adding data about courses and certifications	
Add and customize a short description about yourself	Similar circumstances involving education, courses, and credentials, together with a succinct explanation Automatic chronological sorting is by default deactivated, although this function may be turned on as soon as a new experience is added.
The presence of tips for the correct resume for the user	When entering new data, the user is advised at every stage and in practically every area.
Ability to add and customize skills	Any skill or degree of skill may be added.
The section for adding information about which languages the candidate speaks	This area is offered as a customized section however, you may add more languages and skill levels.
Adding a photo and the ability to hide it	Although you can upload photos and conceal them, this is not a particularly logical process.
Adding link data and contact methods	Only a phone number, email address, and LinkedIn profile can be entered.
Resume templates	There are nine distinct templates available.
Choosing the main color of the resume	There are 4-5 color options for each template.
Adding personalized sections to your resume	The only field added to a custom section is a text field.
Section editing	Complete section editing, including changing the section headings and how they appear in the resume. One has the option to select which side the part will be on.
The ability to preview resumes	Only at the very end, just before exporting, is there an option for a preview.
Confirmation of data deletion actions	Since there are no dialog boxes, entering data might be accidentally deleted by the user.
Saving entered data and created templates	Once the resume has been exported, an account will be made, and the template will be kept.
Program usage plan	The program has a monthly fee of \$10. Exporting to PDF is not free of charge.

The final software on the list of examined resume export programs is **LinkedIn**. The focus of this program is not resumes. This social network enables you to discover individuals who share your interests in a single or a variety of disciplines. It also enables applicants to look for employers and vice versa. The user's resume can, however, be exported as a PDF from the profile page. This is the only free method that has been discovered, which is why it was used in the study.

Table 3. Analysis of **LinkedIn** by criteria

Criteria	Description
Ability to add and edit previous work experience	The profile of the social network user provides information on these fields.
Adding education data	
Adding data about courses and certifications	
Add and customize a short description about yourself	
The section for adding information about which languages the candidate speaks	
The presence of tips for the correct resume for the user	There aren't any pointers for writing a resume.
Ability to add and customize skills	A user account is used to add talents to the resume. There is no set level of proficiency for the talent. As well as being picked at random, the highest skill.
Adding a photo and the ability to hide it	Even though it is on the profile, the photo is not placed on the CV.
Adding link data and contact methods	Contact options are also developed from user data however, there are relatively few of them available.
Resume templates	You may not pick a color and there are no templates to build.
Choosing the main color of the resume	
Adding personalized sections to your resume	Modification of sections is not available
Section editing	
The ability to preview resumes	Although no preview is accessible, the file is created rapidly.
Confirmation of data deletion actions	
Saving entered data and created templates	The entered data is saved in the profile
Program usage plan	Forming a resume is free

V. RESULTS

As a result, I am attaching a photo of the preview of the formed resumes since I was able to export only from LinkedIn.

The screenshot shows a resume for Vasyl Ivanochko, a Middle Java developer. It includes a profile section with a bio, an employment history section listing roles at EPAM and SoftJourn, an education section for a Bachelor's degree at Vasyl Stefanyk Precarpathian National University, and a courses section for a Java online marathon. Skills are listed as Teamwork, Java, OOP, Spring, REST, and Angular. The layout is clean and professional.

Figure 1 – pre-view of resume from resume.io

The screenshot shows a resume for Vasyl Ivanochko with a teal header. It includes a personal section with contact information, a work experience section with roles at EPAM and SoftJourn, an education section for a Bachelor's degree, and a skills section with proficiency levels for Java, OOP, Spring, Angular, and REST. The layout is modern and uses a color scheme of teal and grey.

Figure 2 – pre-view of resume from cvmaker.com.ua

The screenshot shows a generated resume for Vasyl Ivanochko with a dark blue header. It includes a contact section with phone and email, a top skills section listing Java, Core Java, and Scrum, an experience section with roles at EPAM Systems and SoftJourn, and an education section for a Bachelor's degree at Vasyl Stefanyk Precarpathian National University. The layout is simple and professional.

Figure 3 – Generated resume from www.linkedin.com

VI. FUTURE RESEARCH

Future research may be about the norms of standards in different countries. And about whether the programs meet these standards.

Also, it can be investigated in the case of performance and design.

Future research may contain a few more programs and maybe create other criteria.

VII. DISCUSSION

In this research created criteria for evaluation of software implementations for automatic resume generation that can be useful for future research or to evaluate new program candidate for a company.

Also, I have demonstrated an example how to examine programs by criteria on three different cost programs. So we can see that this criteria working with a more widely programs like a resume.io and also, with a little plain programs like a LinkedIn.

VIII. CONCLUSION

I examined and discussed three well-known resume-building programs in my study. I compared them and displayed the rough formation findings. Along with the ones mentioned, I also looked at a few other applications that were comparable, but since they did not differ in any way from the ones mentioned, I did not include them in my research.

As we can see, the programs that are functionally enough and operate reliably and rapidly are quite expensive, while those that are free lack all required functions and inadequately address the demands of the ordinary user.

REFERENCES

- [1] S.M. Dias, A.L. Gancarski and P.R. Henriques, "Automatic Generation of CVs from Online Social Networks," *Communications in Computer and Information Science*, vol. 563, pp 258–263, 2015, doi: 10.1007/978-3-319-27653-3_25
- [2] E. J. Gonzalez et al., "Intelligent Agents and Apache Cocoon for a CV Generation System," *2007 IEEE/ACS International Conference on Computer Systems and Applications*, 2007, pp. 9-15, doi: 10.1109/AICCSA.2007.370858.
- [3] Resume.io. (2022). [Online]. Available: <https://resume.io>
- [4] Cvmaker. (2022). [Online]. Available: <https://www.cvmaker.com.ua>
- [5] LinkedIn. (2022). [Online]. Available: <https://www.linkedin.com>

Overview of methods and existing solutions for task order generation

Yuliia Dudii

*Department of Information Technology
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine*

Abstract—This is a review paper related to the problem of working time planning and its automatic prioritization. The paper presents an overview of theoretical methods and existing facilities, as well as a comparison of properties of existing software implementations according to selected criteria.

Keywords—*Application, Task, Priority, Planning, Evaluation, Comparison, Automation, Pomodoro.*

I. INTRODUCTION

People accumulate obligations around them over time. They frequently underestimate their capacity and assign the wrong priorities to their duties, which results in poor time management. Minor errors become too many significant issues future. These issues eventually impact an individual's mental health and ruin his life.

My study led to an examination of current programs, which would aid in helping individuals schedule their time properly and help them prevent such errors.

II. ANALYSIS OF LATEST RESEARCH AND PUBLICATIONS

This research based on some of other publications. One of them is more psychological than technical. Article [1] is discussed about efficiency of people with their work. It described how their procrastinating and how to solve this problem. This study explained how planned and predefined tasks are more productive in contradistinction to planning tasks on the go.

The publication [2] describes how important planning assistants are in the work. It is important for a person to see his graph in front of him visualized in beautiful graphs and colors. It is crucial on a psychological and a physical level both. I analyzed such assistants in my work.

It is important to maintain a balance between life and work. Take breaks, rest, do not work overtime is best practices. Ideally, planning assistants should take this into account and provide guidance. More detailed about work-life balance are published in work [3].

Some of criteria based on research [4]. Interaction of three core aspects to a personalized scheduling task are described here.

III. EVALUATION CRITERIA OF EXISTING SOLUTIONS

I analyzed and tested selected apps to develop assessment criteria. It will be further explained. These criteria include the following end-user capabilities:

- the possibility of keeping a notebook for personal use;
- the possibility of creating a task for execution;
- the possibility of editing the task;
- the possibility of deleting the task;
- the possibility of adding a description of the task;
- the possibility of giving priority to the task;
- the possibility of adding a date from which it would be convenient to complete the task;
- the possibility of adding a date by which the task must be completed;

- the possibility of adding an estimated duration of the task;
- the possibility of creating a duplicate task;
- the possibility of changing the status of tasks;
- the ability to display completed tasks;
- the possibility of granting the status "Completed" for the task;
- the possibility of granting the status "Scheduled" for the task;
- the possibility of granting the status "Active" for the task;
- the ability to sort tasks by priority;
- the ability to start performing a task that will last a clearly defined period;
- the ability to pause a task that should last a clear period;
- the possibility of a built-in Pomodoro time management system;
- the possibility of automating the selection of the most successful time to perform the task.

IV. ANALYSIS OF COMPETITORS

Three time-management tools have been selected for this analysis. I choose the Todoist as the first tool for comparison among them [5]. Although this tool is relatively simple to use, it is not without flaws.

The application is exactly a notebook for maintaining personal work, not for a team, as most comparable apps are. If you meet the requirements outlined above. You may use it to carry out common work procedures on jobs like creation, editing, and deletion. You can provide an additional description and a priority ranging from 4 to 1 when creating a task. The program's first flaw is that there is only the option to create a task for a certain day. It is not possible to specify a date from which it would be suitable to start the activity. Additionally, neither the option to automatically choose a day for the work nor the opportunity of establishing a deadline for completion of this activity are present. It is important to note that the program's duties are simple to administer, that the user-friendly interface is comfortable and clear, that duplication features are available, and that the status may be changed. The ability to display tasks in this program, both finished and planned, is nicely setup, but sadly, there is no way to change the task's state while it is still in process. Additionally, there is automatic sorting. Another drawback is that it is hard to keep track of how long a task takes.

The Nifty program was the following research initiative [6]. A complicated and loaded software with few functionalities compared to rivals. You need to spend a lot of time in the application to learn to use them. Although there is a distinct component for personal projects, this program is largely focused on collaboration. The typical ability to add, amend, and remove tasks is included. Surprisingly, no feature would let the work and its benefit have a description included. The program offers a variety of choices for managing time and dates, scheduling upcoming tasks, and keeping track of accomplished tasks. This is further demonstrating that this work application is more geared toward teamwork. Strangely, finished tasks don't displayed anywhere and just vanish, no dialog boxes signal their completion. The task status is the same as it was in the previous application however, this program does not allow you to arrange jobs by priority, which is a major drawback. The application also allows you to begin a task that will last for a predetermined period.

The last one for this study was the Worksection program [7]. The program is very rich in functionality. The intriguing aspect is that you can put the first task in this application with values from 1 to 10. It does not contain many disadvantages. Among the key ones - there is a missing function of creating a duplicate. There is no built-in Pomodoro time management technique just like in other programs. These techniques would allow you to increase your productivity by reducing the influence of external and internal distractions.

V. RESULTS

I summarized the results in a table that quickly describes whether this or other functional possibilities are present or not to make the data easier to use.

Table 1 - Summary table of the presence of functionality in the studied implementations according to the criteria

№	Criteria	Todoist	Nifty	Worksection
1.	The possibility of keeping a notebook for personal use	+	+	+
2.	Ability to create a task to be performed	+	+	+
3.	Ability to edit the task	+	+	+
4.	Ability to delete the task	+	+	+
5.	The possibility of adding a description of the task	+	–	+
6.	Ability to prioritize a task	+	–	+
7.	The possibility of adding a date when it would be convenient to complete the task	–	+	+
8.	Ability to add a date by which the task must be completed	+	+	+
9.	The possibility of adding an estimated duration of the task	–	–	+
10.	Ability to duplicate a task	+	+	–
11.	Ability to change the task status	+	+	+
12.	Ability to display completed tasks	+	–	+
13.	Ability to provide the status "Completed" for the task	+	+	+
14.	Ability to provide the status "Planned" for the task	+	+	+
15.	Ability to provide the status "Active" for the task	–	–	+
16.	Ability to sort tasks by priority	+	–	+
17.	The ability to start a task that will last for a clearly defined period	–	+	+
18.	Ability to pause a task that should last for a specified period	–	–	–
19.	Possibility of built-in Pomodoro time management technique system	–	–	–
20.	The possibility of automating the selection of the most successful time to perform the task	–	–	–

The Worksection application offers the most advantages and hence the fewest drawbacks, according to the as outlined below table (16 advantages, 4 disadvantages). Among the examined applications, the Todoist program comes in second (pluses: 13, minuses: 7), and Nifty comes in last (pluses - 11, minuses - 9).

It is important to note that none of the examined applications included the features of "automating the selection of the optimal time to accomplish the task" or "built-in system of Pomodoro time management approach.". Therefore, these tools make it easier for the user to manage his duties and evaluate his time, but they do not do it automatically for the user.

VI. FUTURE RESEARCH

Future research might concentrate additionally attention on studies of fault tolerance, uptime, and performance of diverse applications. It is also feasible to divide this topic's mobile application investigation into different research. Future works can analyze how well these programs adhere to the norms and standards might be crucial.

VII. CONCLUSION

I thoroughly examined several already existing solutions, including Nifty, Todoist, and Worksection. I analyzed the requirements to develop an application for generating the order of execution of planned tasks. I concluded that the current solutions still require some improvements because each has specific drawbacks. The end user may find it to be very significant.

No app implements built-in progressive methods of time management or some optimization of work.

REFERENCES

- [1] D. Ariely and K. Wertenbroch, "Procrastination, deadlines, and performance: Self-control by precommitment", *Psychological Science*, vol. 13, no. 3, pp. 219–224, 2002, doi: 10.1111/1467-9280.00441
- [2] J. Bank, Z. Cain, Y. Shoham, C. Suen, and D. Ariely, "Turning personal calendars into scheduling assistants", *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, 2012, doi: 10.1145/2212776.2223854
- [3] N. Yusop, "A systematic literature review: Optimization timetable in education to support work-life balance (WLB)", *Journal of Computing Research and Innovation*, vol. 7, no. 2, pp. 316–326, 2022, doi: 10.24191/jcrinn.v7i2.324
- [4] Pauline Berry, Melinda Gervasio, Bart Peintner, and Neil Yorke-Smith, "Balancing the Needs of Personalization and Reasoning in a User-Centric Scheduling Assistant", Technical Note No. 561, Feb. 2007
- [5] *Todoist*. (2007). Doist Inc. Accessed: Nov. 27, 2022. [Online]. Available: <https://todoist.com/>
- [6] *Nifty*. (2019). Nifty Technologies, Inc. Accessed: Nov. 27, 2022. [Online]. Available: <https://nifty.pm/>
- [7] *Worksection*. (2008). [Online]. Available: <https://worksection.com/>

Object Detection in Mobile Robotics

Valentyna Yakovyn, Taras Kriukov

*Department of Information Technology
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine*

Abstract—*This paper presents an improved object detection method based on a combination of one-stage object detector followed by an additional classification stage. The article reports an increase in the IoU metric by 8%, classification accuracy by 7%, and f1-score by 5%. The research was conducted using computational experiments and modeling techniques. The results were discussed and interpreted.*

Keywords—*computer vision, object detection, image classification, machine learning, mobile robotics, software tools*

I. INTRODUCTION

The use of Artificial Intelligence (AI) is a great example of advances, which helps to detect various objects more accurately, which in some cases is important, for example, to identify objects on the road. The Object Detection (OD) is a computer technology whose main purpose is to recognize objects by processing images and highlight them with bounding boxes according to a certain classification. Unusual photos (image size, different viewpoints, mixing the object with the background), real-time detection rate, limited data, and class imbalance can cause significant problems for OD. Our research was triggered by the impressions of the use cases of artificial intelligence in various applications, presented at scientific conferences and challenges. The main purpose of this work is to develop a reliable object detector to be used in mobile robotics. It should detect other mobile robots and distinct them from other (non-robot) objects.

II. RELATED WORK

Today, many classical OD algorithms are used in various fields of human activity. An approach to knowledge-assisted semantic video object detection based on a multimedia ontology infrastructure is presented in [1]. Paper [2] presents an approach to activity recognition based on detecting and analyzing the sequence of objects that are being manipulated by the user. In domains such as cooking, where many activities involve similar actions, object-use information can be a valuable cue. In order for this approach to scale to many activities and objects, however, it is necessary to minimize the amount of human-labeled data that is required for modeling. Paper [2] describes a method for automatically acquiring object models from video without any explicit human supervision. It was shown in the [3] that the You Only Look Once (YOLO) algorithm, particularly, YOLOv4, object detection neural network based on the CSP approach, scales both up and down and is applicable to small and large networks while maintaining optimal speed and accuracy. They proposed in [3] a network scaling approach that modifies not only the depth, width, resolution, but also structure of the network. In [4], the authors proposed a novel object detection network by considering the difference between the localization and detection tasks. Algorithms such as YOLO [5], Fast R-CNN [6], HOG, and Mask R-CNN [7] are considered among the best. Each of them has their own advantages, for example YOLO is considered one of the best OD algorithms due to its speed. The processing speed of the image in real time allows processing of many frames per second and newer modifications of the algorithm can process more than hundred frames per second. Paper [8] presents various Object Detection Algorithms such as face detection, skin detection, color detection, shape detection, target detection are simulated and implemented to detect various types of objects with improved accuracy.

III. METHODOLOGY

Participating in the specialized scientific competition "Roborace" is a great way to test the efficiency and accuracy of the neural network, because due to unpredictable movement of the robot camera, we get a blurred image, which can cause problems with recognition and classification of the object. The technical challenge of the competition is to build an autonomous robot [9]-[12]. This includes the resolving of the following problems: in a shortest time the autonomous robot has to pass the set number of laps from the starting position to the finish, avoiding other robots and without going outside the track. The main data input device is the camera. Therefore, we proposed to use the additional classification stage after the OD stage. We can filter out false positives by additional image classification in the constraint frame and selecting a confidence threshold. All the models have been implemented with the open-source library TensorFlow 2. This framework is designed specifically for deep machine learning. Keras was used as a high-level, user friendly framework.

IV. RESULTS

The result of our work is the coordinate detection with an IoU metric increased by 8% in comparison to detection without extra classification. Also, Fig. 1 and Fig. 2 present the classification metrics.

```
[8]: cm = confusion_matrix(y_true, y_pred)
      print(cm)
      print(classification_report(y_true, y_pred, target_names=target_names))
```

	[[472 140]				
	[80 235]]				
		precision	recall	f1-score	support
Not a Robot		0.86	0.77	0.81	612
Robot		0.63	0.75	0.68	315
accuracy				0.76	927
macro avg		0.74	0.76	0.75	927
weighted avg		0.78	0.76	0.77	927

Fig. 1. Classification report for YOLO detection

```
[40]: cm = confusion_matrix(y_true, y_pred)
       print(cm)
       print(classification_report(y_true, y_pred, target_names=target_names))
```

	[[524 88]				
	[77 238]]				
		precision	recall	f1-score	support
Not a Robot		0.87	0.86	0.86	612
Robot		0.73	0.76	0.74	315
accuracy				0.82	927
macro avg		0.80	0.81	0.80	927
weighted avg		0.82	0.82	0.82	927

Fig. 2. Classification report for YOLO detection with extra classification stage

V. DISCUSSION

We achieved this result by using a classifier head after OD, which helped reduce the number of false positives. It should be noted that our neural network does not always accurately recognize the selected objects; the reason for this is that at the stage of recognition the network may receive a very indistinct image and the system may not correctly identify the object in the image. Also, the use of our method is limited by the computational hardware of a mobile robot and power and capacity of a power supply.

VI. FUTURE RESEARCH

In the future, the direction of research will be aimed at improving the detection quality and reducing the training time and computational complexity of the model.

VII. CONCLUSION

In this paper we proposed the enhanced method for OD for mobile robotics. It is based on the stack of OD and classification heads implemented using artificial neural networks. It provides an increase in the IoU metric by 8%, classification accuracy by 7%, f1-score by 5%.

VIII. ACKNOWLEDGMENT

We are grateful to the researchers of the Department of Information Technology of Vasyl Stefanyk Precarpathian National University for the professional advice provided to us in the research process.

IX. DISCLOSURES

The authors have nothing to declare.

REFERENCES

- [1] S. Dasiopoulou, V. Mezaris, I. Kompatsiaris, V.-K. Papastathis, and M. G. Strintzis, "Knowledge-assisted semantic video object detection," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 10, pp. 1210-1224, Oct. 2005, doi: 10.1109/TCSVT.2005.854238
- [2] J. Wu, A. Osuntogun, T. Choudhury, M. Philipose, and J. M. Rehg, "A Scalable Approach to Activity Recognition based on Object Use," *2007 IEEE 11th International Conference on Computer Vision*, 2007, pp. 1-8, doi: 10.1109/ICCV.2007.4408865
- [3] C -Y. Wang, A. Bochkovskiy, and H -Y. M. Liao, "Scaled-YOLOv4: Scaling Cross Stage Partial Network," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 13024-13033, doi: 10.1109/CVPR46437.2021.01283
- [4] J. U. Kim and Y. Man Ro, "Attentive Layer Separation for Object Classification and Object Localization in Object Detection," *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 3995-3999, doi: 10.1109/ICIP.2019.8803439
- [5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91
- [6] R. Girshick, "Fast R-CNN," *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169
- [7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2980-2988, doi: 10.1109/ICCV.2017.322
- [8] A. Raghunandan, Mohana, P. Raghav, and H. V. R. Aradhya, "Object Detection Algorithms for Video Surveillance Applications," *2018 International Conference on Communication and Signal Processing (ICCSP)*, 2018, pp. 0563-0568, doi: 10.1109/ICCSP.2018.8524461
- [9] М. Козленко, "Система компенсації гравітаційного впливу на лінійну швидкість автономного мобільного колісного робота," *І Науковий семінар Кіберфізичні системи досягнення та виклики*, 2015, с. 149-157
- [10] М. Козленко, "Модуль інерціальної навігації автономного мобільного робота та аналіз характеристик його сигналів," *Вісник Хмельницького національного університету Технічні науки*, № 6, с. 160-164, 2014
- [11] М. Козленко, "Характер та симетрія функцій правдоподібності при демодуляції шумоподібних сигналів в мобільній робототехніці," *Вісник Хмельницького національного університету Технічні науки*, № 6, с. 135-139, 2013
- [12] М. Kozlenko, "The interference immunity of the telemetric information data exchange with autonomous mobile robots," *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, no. 1, pp. 107–113, 2015

Supervised Machine Learning Based Signal Demodulation in Chaotic Communications

Mykola Kozlenko

*Department of Information Technology
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine*

 <https://orcid.org/0000-0002-2502-2447>

Abstract—A chaotic modulation scheme is an efficient wideband communication method. It utilizes the deterministic chaos to generate pseudo-random carriers. Chaotic bifurcation parameter modulation is one of the well-known and widely-used techniques. This paper presents the machine learning based demodulation approach for the bifurcation parameter keying. It presents the structure of a convolutional neural network as well as performance metrics values for signals generated with the chaotic logistic map. The paper provides an assessment of the overall accuracy for binary signals. It reports the accuracy value of 0.88 for the bifurcation parameter deviation of 1.34% in the presence of additive white Gaussian noise at the normalized signal-to-noise ratio value of 20 dB for balanced dataset.

Keywords—deterministic chaos, chaotic signal, chaotic communications, bifurcation, bifurcation parameter keying, demodulation, bit error rate, machine learning, deep learning, convolutional neural network

I. INTRODUCTION

Deterministic chaotic signals reflect the almost unpredictable dynamical behavior based on simple nonlinear rule [1]. Chaotic signals are successfully used as carriers in analog and digital communication systems [2], [3]. Chaotic modulation is an efficient wideband keying scheme [4]. It performs as a spread spectrum telecommunication system. Ergodic Chaotic Parameter Modulation (ECPM) is one of such keying schemes [5]. A modulation scheme determines how bits are mapped to the bifurcation parameter [6] of transmitted signals [7]. It requires sophisticated demodulation algorithms that increase the demodulator complexity. A logistic map [8] is a simple rule that can generate chaotic behavior. Logistic map based applications are widely used in various fields like digital communications, cybersecurity, etc. [9], [10]. This paper presents a bifurcation parameter demodulation approach for signals generated using the logistic map. It is based on supervised machine learning and considers communication in the presence of additive white Gaussian noise.

II. RELATED WORK

Chaos has been applied in communication since 90s [11]. Paper [12] presents an overview of chaotic communications and basic chaotic modulation schemes like Chaos Shift Keying (CSK), Differential Chaos Shift Keying (DSK), Additive Chaos Modulation (ACM), and Multiplicative Chaos Modulation (MCM). It also describes synchronized chaotic systems and direct chaotic communication. Reference [13] proposes an information security inversion technology based on time series. It includes chaotic signal demodulation technology, chaotic channel modeling, and chaotic signal interference channel problem analysis. Paper [14] presents the particle filter to demodulate the Chaotic Parameter Modulation (CPM) signal. In [15] the state equations of unscented particle filter (UPF) for the demodulation of chaotic parameter have been derived. A novel modified algorithm based on UPF for the demodulation has been presented [15]. The papers [16] – [18] solve the problem of weak radio signals demodulation for Frequency Shift Keying (FSK) and Phase Shift Keying (PSK) signals with deep Machine Learning (ML) using

convolutional neural networks (CNN) as well as fully connected networks. Paper [19] addresses the demodulating of underwater communication signal problem as a classification task and conceives a machine learning based demodulation scheme. Reference [20] presents the security weakness of communication method based on parameter modulation of a chaotic system and adaptive observer-based synchronization scheme. It has been shown that the security can be compromised without exact knowledge of the chaotic system properties [20]. Paper [21] presents the idea of using CNNs with deep learning structure to predict future symbols based on the received signal, to further reduce inter-symbol interference and to obtain a better Bit Error Rate (BER) performance in chaotic baseband wireless communication systems.

III. METHODOLOGY

Replacing the logistic equation

$$\frac{dx}{dt} = rx(1-x) \quad (1)$$

with the quadratic recurrence equation

$$x_{n+1} = rx_n(1-x_n), \quad (2)$$

where r is a positive constant gives the logistic map [22]. Chaotic signals in this research are generated with the logistic map using Eq. 2. Fig. 1 presents the bifurcation diagram for r in the range from 2.8 to 4.0.

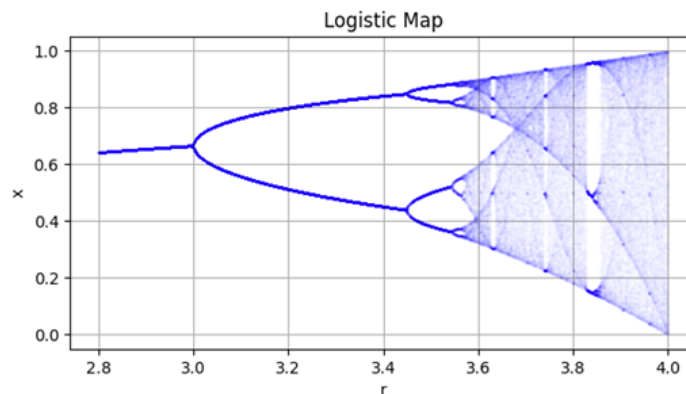


Fig. 1. Bifurcation diagram for the logistic map

Fig. 2 shows the time domain waveforms of the chaotic signals.

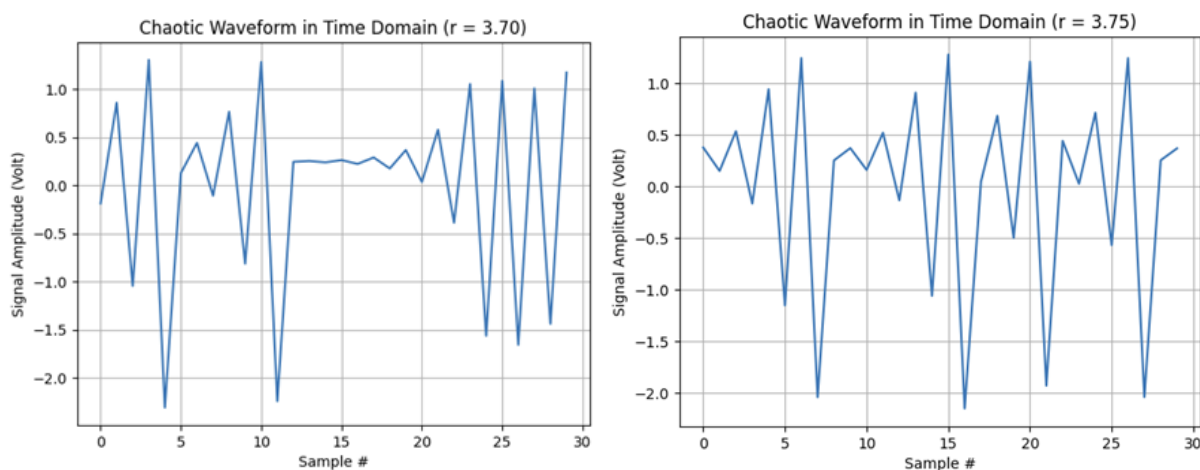


Fig. 2. Time domain waveforms of chaotic signals for parameter $r = 3.7$ and $r = 3.75$

In this research, the modulation scheme represents digital data as a variation in the bifurcation parameter of a chaotic baseband carrier wave. The deviation of the bifurcation parameter value is 0.05; from 3.7 for "space" to 3.75 for "mark" (percent deviation equals 1.34%). The baseband bandwidth is 5512.5 Hz. The sampling frequency is 11025 Hz. The bit duration time is 0.3715 s, 4096 samples (exactly as it is in JT65 protocol). The bit rate is 2.67 bit/s. The signal-to-noise ratio (SNR) is -13 dB, the normalized SNR (E_b/N_0) is +20 dB, and the spreading factor is +33 dB).

Fig. 3 shows the structure of the convolutional neural network that is used for the demodulation.

```

Model: "sequential"
Layer (type)                Output Shape                Param #
-----
reshape (Reshape)           (None, 4096, 1)            0
batch_normalization (Batch  (None, 4096, 1)            4
conv1d (Conv1D)             (None, 4096, 128)         2176
batch_normalization_1 (Batch (None, 4096, 128)         512
flatten (Flatten)           (None, 524288)             0
dense (Dense)               (None, 64)                 33554496
batch_normalization_2 (Batch (None, 64)                 256
dense_1 (Dense)             (None, 2)                  130
-----
Total params: 33,557,574
Trainable params: 33,557,188
Non-trainable params: 386

```

Fig. 3. Structure of the convolutional neural network

The size of the training dataset is 12800 records, the validation set size is 3200, and the test set size is 4000 records. Loss function: categorical cross-entropy, optimizer: adam. Two output neurons are used in the model and labels are one-hot encoded. Thus, the model can be easily converted to a multiclass classifier for m-ary signals.

IV. RESULTS

The classification report for the test data is shown in Table. 1.

TABLE 1. CLASSIFICATION REPORT

Class	Classification Metrics			
	Precision	Recall	F1-Score	Support
0 – "space"	0.94	0.82	0.87	2003
1 – "mark"	0.84	0.95	0.89	1997
Accuracy			0.88	4000
Macro avg	0.89	0.88	0.88	4000
Weighted avg	0.89	0.88	0.88	4000

V. DISCUSSION

The purpose of this study was to gain a better understanding of the ability of machine learning algorithms to demodulate and decode the baseband chaotic signals. The result of the present study supports the hypothesis that chaotic signal generated with the logistic map can be demodulated with ML techniques. The results of this research provide supporting evidence that it is possible even in the presence of noise. This is the main take away from this paper. This pattern of results is consistent with the previous works [16], [18] those deal with ML-based demodulation for frequency shift keying modulation scheme. These results represent the direct demonstration of chaotic signal demodulation within the scope of the ML approach. There are at least three potential limitations concerning the results of this study. A first limitation is that we used only signals generated with the logistic map. A second potential limitation is that we used supervised machine learning approach. This means that the algorithm should be trained with previously known signal patterns. Unsupervised approach looks like more promising technique. The third limitation is that we considered only the case of AWGN. Despite these limitations, the results suggest practical implication that chaotic signals can be demodulated without exact knowledge of the bifurcation parameter values.

VI. FUTURE RESEARCH

In terms of future research, it would be useful to extend the current findings by examining other types of chaotic maps. Also, it would be useful to research the ability of unsupervised ML approach in the context of this task. Other types of interferences should be studied as well.

VII. CONCLUSION

The main conclusion that can be drawn is that chaotic baseband signals generated with the logistic map in the bifurcation parameter modulation scheme can be demodulated using supervised machine learning approach. In summary, this paper presents the values of quality metrics. The overall accuracy is 0.88 at normalized SNR value of +20 dB.

VIII. ACKNOWLEDGMENT

The author gratefully acknowledges the contributions of scientists of the Department of Information Technology of the Vasyl Stefanyk Precarpathian National University for scientific guidance given in discussions and technical assistance helped in the actual research.

IX. ETHICS DECLARATIONS

The author has nothing to disclose.

REFERENCES

- [1] R. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, pp. 459–467, 1976, doi: 10.1038/261459A0
- [2] C. Wang, N. Jiang, G. Li, C. Xue, A. Zhao and K. Qiu, "Secure chaotic communication with spectrum expansion/compression," *2017 16th International Conference on Optical Communications and Networks (ICOON)*, 2017, pp. 1-3, doi: 10.1109/ICOON.2017.8121333
- [3] G. Tang, L. Zhu, Q. Wu, Q. He and L. Yu, "A hybrid spread spectrum communication method based on chaotic sequence," *2021 International Symposium on Networks, Computers and Communications (ISNCC)*, 2021, pp. 1-5, doi: 10.1109/ISNCC52172.2021.9615817
- [4] N. X. Quyen, V. Van Yem and T. M. Hoang, "Chaotic modulation based on the combination of CPPM and CPWM," *Proceedings of the Joint INDS'11 & ISTET'11*, 2011, pp. 1-6, doi: 10.1109/INDS.2011.6024801
- [5] H. Leung, H. Yu, and K. Murali, "Ergodic chaos-based communication schemes," in *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 66, no. 3 Pt 2A, 2002, pp. 036203, doi:10.1103/PhysRevE.66.036203
- [6] J. Borwein, and D. Bailey, "Bifurcation points in the logistic iteration," in *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, Wellesley, MA: A K Peters, pp. 50-53, 2003

- [7] H. N. Abdullah and I. F. Ali, "Design of chaotic bifurcation parameter modulation scheme using mean value estimation," *Iraqi Journal of ICT*, vol. 1, no. 1, pp. 52–61, May 2018, doi: 10.31987/ijict.1.1.13
- [8] T. Tsuchiya and D. Yamagishi, "The complete bifurcation diagram for the logistic map," *Zeitschrift für Naturforschung A*, vol. 52, no. 6-7, 1997, pp. 513-516, doi: 10.1515/zna-1997-6-708
- [9] M. M. de A. Kotaki and M. Luppe, "FPGA implementation of a pseudorandom number generator based on k – logistic map," *2020 IEEE 11th Latin American Symposium on Circuits & Systems (LASCAS)*, 2020, pp. 1-4, doi: 10.1109/LASCAS45839.2020.9068999
- [10] Z. Mu and H. Liu, "Research on digital media image encryption algorithm based on Logistic chaotic map," *2020 International Conference on Robots & Intelligent System (ICRIS)*, 2020, pp. 108-111, doi: 10.1109/ICRIS52159.2020.00035
- [11] S. Hayes, C. Grebogi, E. Ott, "Communicating with chaos," *Phys Rev Lett*, vol. 70, no. 20, 1993, pp. 3031-3034, doi:10.1103/PhysRevLett.70.3031
- [12] A. Riaz and M. Ali, "Chaotic communications, their applications and advantages over traditional methods of communication," *2008 6th International Symposium on Communication Systems, Networks and Digital Signal Processing*, 2008, pp. 21-24, doi: 10.1109/CSNDSP.2008.4610808
- [13] Z. Xing, "Chaotic wireless communication and related network information security inversion technology," *2020 International Conference on Advance in Ambient Computing and Intelligence (ICAACI)*, 2020, pp. 36-41, doi: 10.1109/ICAACI50733.2020.00012
- [14] M. Xu, Y. Song and L. Liu, "A novel demodulation method for chaotic parameter modulation," *2006 International Conference on Wireless Communications, Networking and Mobile Computing*, 2006, pp. 1-4, doi: 10.1109/WiCOM.2006.156
- [15] H. Jin-feng and G. Jing-bo, "A novel UPF-based demodulator for chaotic parameter modulation systems," *2007 2nd IEEE Conference on Industrial Electronics and Applications*, 2007, pp. 2575-2578, doi: 10.1109/ICIEA.2007.4318878
- [16] M. Kozlenko and V. Vialkova, "Software defined demodulation of multiple frequency shift keying with dense neural network for weak signal communications," *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2020, pp. 590-595, doi: 10.1109/TCSET49122.2020.235501
- [17] I. Lazarovych *et al.*, "Software implemented enhanced efficiency BPSK demodulator based on perceptron model with randomization," *2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2021, pp. 221-225, doi: 10.1109/UKRCON53503.2021.9575458
- [18] M. Kozlenko, I. Lazarovych, V. Tkachuk, and V. Vialkova, "Software demodulation of weak radio signals using convolutional neural network," *2020 IEEE 7th International Conference on Energy Smart Systems (ESS)*, 2020, pp. 339-342, doi: 10.1109/ESS50319.2020.9160035
- [19] C. Li, F. Marzani, and F. Yang, "Demodulation of chaos phase modulation spread spectrum signals using machine learning methods and its evaluation for underwater acoustic communication," *Sensors*, vol. 18, no. 12, p. 4217, Dec. 2018, doi: 10.3390/s18124217
- [20] G. Álvarez, F. Montoya, M. Romera, and G. Pastor, "Breaking parameter modulated chaotic secure communication system," *Chaos, Solitons & Fractals*, vol. 21, no. 4, 2004, pp. 783-787, doi: 10.1016/j.chaos.2003.12.041
- [21] Hai-Peng Ren, Hui-Ping Yin, Hong-Er Zhao, C. Bai, C. Grebogi, "Artificial intelligence enhances the performance of chaotic baseband wireless communication," *IET Communication*, vol. 15, 2021, pp. 1467–1479, doi: 10.1049/cmu2.12162
- [22] Wolfram MathWorld, "Logistic map," *Wolfram MathWorld*. [Online]. Available: <https://mathworld.wolfram.com/LogisticMap.html> [Accessed Nov. 28, 2022]

Scientific electronic edition
Наукове електронне видання

2022 International Conference on Innovative Solutions in Software Engineering (ICISSE)

Conference Proceedings

November 29-30, 2022
Ivano-Frankivsk, Ukraine

Issue editors: Mykola Kuz and Mykola Kozlenko

Publication language: English
(includes papers in English and Ukrainian)

First Online: Nov. 29, 2022

Publisher:

Vasyl Stefanyk Precarpathian National University
1 Stepana Bandery Street
Ivano-Frankivsk, 76018, Ukraine
E-mail: vdvcit@pnu.edu.ua
Tel: +380 (342) 75-13-08

Publisher certificate ДК #7616 2022-05-26

Видавець

Прикарпатський національний університет
імені Василя Стефаника
76018, м. Івано-Франківськ,
вул. С. Бандери, 1, тел.: +380 (342) 75-13-08
E-mail: vdvcit@pnu.edu.ua

Свідоцтво суб'єкта видавничої справи ДК № 7616 від 26.05.2022

ISBN 978-966-640-534-3